2018

# A Novel Method to Analyze DNA Breaks and Repair in Human Cells

Caitlin Elizabeth Goodman
*Wright State University*

# A NOVEL METHOD TO ANALYZE DNA BREAKS AND REPAIR IN HUMAN CELLS

A thesis submitted in partial fulfillment of the requirement for the degree of Master of Science

By

CAITLIN ELIZABETH GOODMAN

B.S., Otterbein University, 2014

2018
Wright State University

WRIGHT STATE UNIVERSITY

GRADUATE SCHOOL

April 09, 2018

I HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER MY SUPERVISION BY Caitlin Elizabeth Goodman ENTITLED A Novel Method to Analyze DNA Breaks and Repair in Human Cells BE ACCEPTED IN PARTIAL FULFLILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF Master of Science

_____

Michael Leffak, Ph.D.
Thesis Director

_____

Madhavi Kadakia, Ph.D
Chair, Department of Biochemistry
and Molecular Biology

Committee on Final Examination

_____
Michael Leffak, Ph.D.

_____
Michael Markey, Ph.D.

_____
Weiwen Long, Ph.D.

_____

Barry Milligan, Ph.D.
Professor and Interim Dean of the
Graduate School

ABSTRACT

Goodman, Caitlin Elizabeth. M.S. Department of Biochemistry and Molecular Biology, Wright State University, 2018. A Novel Method to Analyze DNA Breaks and Repair in Human Cells

Microsatellites repeat sequences are prone to forming non-canonical DNA structures and mutations. These areas of the genome can undergo expansions and contractions and are responsible for a variety of inherited neurological and neuromuscular disorders. Hairpin structures formed by trinucleotide repeats can lead to replication fork stalling, and fork collapse causing DNA double strand breaks. Various mechanisms are involved in processing microsatellites including mismatch repair, base excision repair, and crossover junction endonuclease cleavage. These processes, which are supposed to protect the genome, could also be the culprits which are causing mutations. In order to test and study this hypothesis, the use of a two color marker gene assay to detect DNA double strand breaks at trinucleotide repeats, was used to detect replication fork stalling, and collapse in presence or absence of replication stress. An important mechanism for the restart of a stalled replication fork involves crossover junction endonucleases, which cleave obstacles that prevent passage of the replication fork. This process is led by MUS81 and its associates EME1 and EME2, which form complexes to process these secondary structures allowing the replication fork to progress. My results indicate distinct roles for MUS81-EME1 and MUS81-EME2 complexes in the maintenance of genome stability.

## TABLE OF CONTENTS

## List of Figures

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to take this opportunity to thank all the amazing people in my life. First and foremost, I would like to take this opportunity to thank Dr. Leffak who has been such an amazing mentor throughout this whole experience. His encouragement and guidance have helped me to grow and to learn more than I ever thought possible. The excitement he showed during lab meetings or whenever data was being presented always encouraged me to work with more enthusiasm and to work harder to understand the results and what should come next. I cannot thank him enough for giving me this opportunity to work with him, it was such an invaluable experience that I will always treasure. I would also like to thank my committee members Dr. Long, and Dr. Markey for being a part of my committee. I would also like to thank them for their support, input, and suggestions were irreplaceable. I would also like to thank present and past lab members for all of your suggestions and help. You made the lab such a fun place to be and made it a pleasure to come to school every day; Dr. Joanna Barthelemy, Dr. Eric Romer, Dr. Todd Lewis, Dr. David Hitch, Dr. Sumeet Poudel, Dr. Dean Rider, Rujuta Gadgil, French Damewood IV, and Donovan Christman. I would like to give a special thanks to Dr. Joanna Barthelemy for teaching me everything in the lab and for all her insights. I would also like to give a special thanks to Rujuta Gadgil for helping me with several assays and always being reassuring. Thank you for making this fun crazy experience worthwhile and always being so supportive.

Dedicated to my wonderful husband, and my amazing parents

Charles Goodman

Peg Castagno

Dan Castagno

I could not have done this without your faith, support and constant encouragement. Thank you for always being a source of inspiration and enthusiasm. You all mean the world to me.

# INTRODUCTION

## Microsatellite Instability

Microsatellites are short repetitive sequences of DNA 1-6 bp long, can be repeated 5-50 times, and are found everywhere in the genome. These short tandem repeats make up around 3% of the entire genome and are areas that are highly prone to mutations which can lead to disease (Reid et al. 1998, Gadgil et al. 2016). There are over thirty human neurological and neuromuscular inherited disorders associated with microsatellite instability (Figure 1) (Kim, Mirkin 2013). Most of these disorders are caused by the expansion of trinucleotide repeats (CGG)•(CCG), (CAG)•(CTG),(GAA)•(TTC), and (GCN)•(NGC) however, diseases also occur at other microsatellite loci. In most individuals, repeat tracts are shorter than ~30 repeats and are generally stable; however in longer repeats, expansions and mutations lead to earlier and enhanced disease progression and severity of disease (Schmidt et al. 2016, Mirkin et al. 2007).

Various downstream effects occur depending on where the microsatellite expansion is located in the genome. In Figure 1, if expansions occur in the exon region it can cause a shift of the reading frame, or over representation of a repeated amino acid. Within the intron or untranslated regions microsatellite expansion can lead to improper splicing and faulty transcript formation (Cummings, et al. 2000).

Healthy individuals typically display 30 or fewer repeats per microsatellite, while tracts of 50 or more repeated units are associated with disease states, as seen in Figure 1. (Liu et al. 2012). Myotonic Dystrophy Type 1 or DM1 affects about 1 in every 8000 people. It is an autosomal dominant disease, which results from CTG repeat expansion in the 3'-untranslated region of DMPK gene (Cummings et al. 2000). This produces a toxic-gain-of-function CUG RNA.

Within microsatellites, once the ~50 repeat threshold is reached, the microsatellites tend to expand. This bias is passed to each generation such that the CTG tract becomes progressively larger, with earlier onset during life, and can reach up to 4000 repeats. It has been observed that the severity of the disease is often correlated with the length of the repeat tract. The predisposition of these unstable tracts to expand with each generation within a family provides a molecular explanation for genetic anticipation (Wells et al. 1998, Seriola et al. 2011).

The complete mechanism behind microsatellite repeat instability is not yet known. It appears that the CTG tract of DM1 patients shows both somatic and intergenerational instability. This seems to occur by one of two mechanisms. At around 50-80 repeats, or the lower end of the disease range, repeat expansion occurs as a function of length. This causes repeats with relatively large numbers to become more frequent along with the increasing repeat number. After CTG repeats exceed a threshold of 80 repeats the pattern of instability changes

drastically, which results in a high frequency of distinct mutational expansions leading to progressively longer repeats. (Seriola et al. 2011). These mutational expansions can occur either through numerous small scale expansions, or through one large scale expansion event. The frequent small scale events occur through mutations such as strand slippage followed by hairpin formation on the nascent DNA strand which can lead to small scale expansions. This process repeated through multiple replication events leads to progressively larger repeats. Larger repeats, where a microsatellite can more than double in a single replication event, may occur through break induced replication. This replication mechanism could explain how severe DM1 diseases cases could have up to 4,000 repeats of $(CTG)_n$ in the 3' untranslated region of DMPK (Kim et al. 2017). These large expansions are the basis behind the genetic anticipation hypothesis seen in myotonic dystrophy, Huntington's disease, and other microsatellite repeat diseases. This, along with data from numerous other studies has led to the theory that non-canonical DNA and/ or double strand breaks in microsatellites are essential in CTG expansions (Seriola et al. 2011).

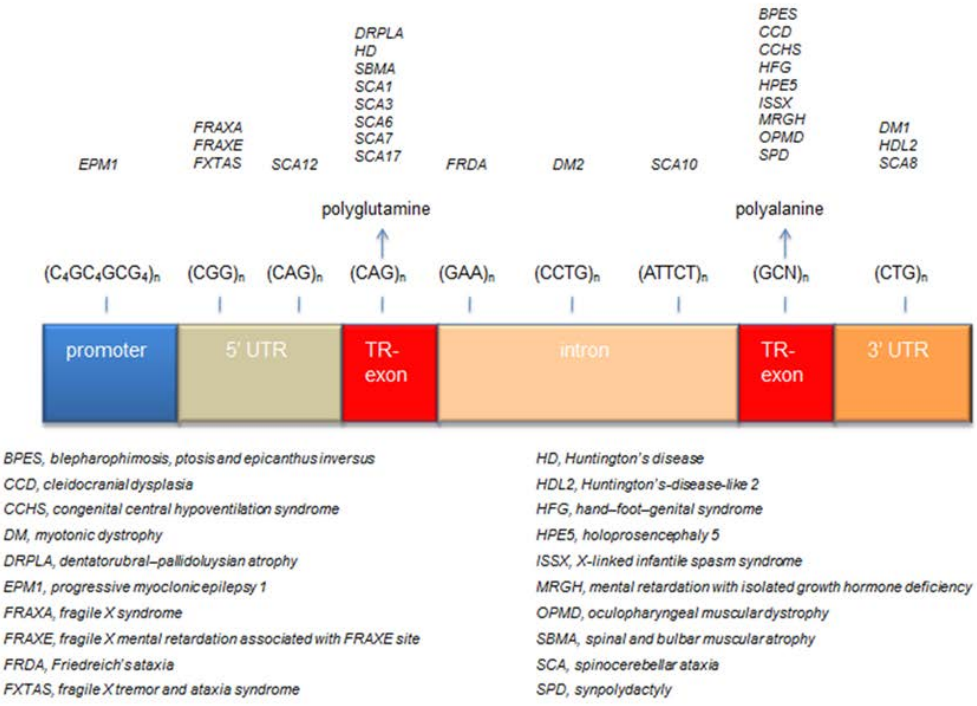



Figure 1. **Unstable microsatellites cause human disease.** Various diseases caused by microsatellite repeat expansion and locations of these repeats within the genome.

**Non-B DNA leads to expansion or contraction**

Microsatellites are abundant within the genome and have the tendency to form various non-canonical (i.e. non-B) DNA structures. The structures produced are the main causative factors in genetic instability and human disease. These structures include hairpins, cruciform, triplex or H-DNA, G4 quadruplex, slipped DNA, and sticky DNA (Figure 2) (Mirkin 2006). Trinucleotide repeats, whose expansion is involved in over a dozen hereditary neurological disorders, support the mutagenic role of non-canonical DNA structures. One of the non-B DNA structures formed is cruciforms or hairpins. These form at repeats such as (CAG)•(CTG) when the duplex is unwound during replication. This allows the repeat to base pair with itself forming a hairpin structure. These structures have been found to induce DNA double strand breaks (DSBs), deletions and translocations (Zhao et al. 2010, Wang et al. 2009).

The tendency of trinucleotide repeats to undergo expansions or contractions is based on whether the non B-DNA structures are present in the nascent or template strand of the DNA (Mirkin et al. 2007). The predisposition of (CTG)•(CAG) repeats to cause DNA instability correlates directly with their formation of hairpins (Figure 3). Models of TNR instability show hairpin formation during DNA synthesis either during cell division or as a stage of DNA repair. It has also been shown that DM1 patient cells grown in culture with replication stressing agents such as aphidicolin, or hydroxyurea increased (CTG)•(CAG) instability at the DMPK locus (Liu et al. 2010).

Healthy length trinucleotide repeats are more inclined to contract rather than expand during replication. This is needed to maintain normal alleles and to maintain genetic stability. The longer the microsatellite, the more stable the hairpin structure. As the hairpin structure becomes more stable, the template tends toward instability and expansions, leading to an avalanche of events with the tendency toward increased expansion (Pelletier et al 2005). Contributing factors of small scale expansion and contractions can be attributed to polymerase slippage as well as environmental stress factors (Gacy et al. 1995, Chatterjee et al. 2015). Expansion of DNA has only been observed with structure forming repeat sequences leading to the conclusion that non-B DNA structures are needed in order for expansion to take place (Gacy et al. 1995).

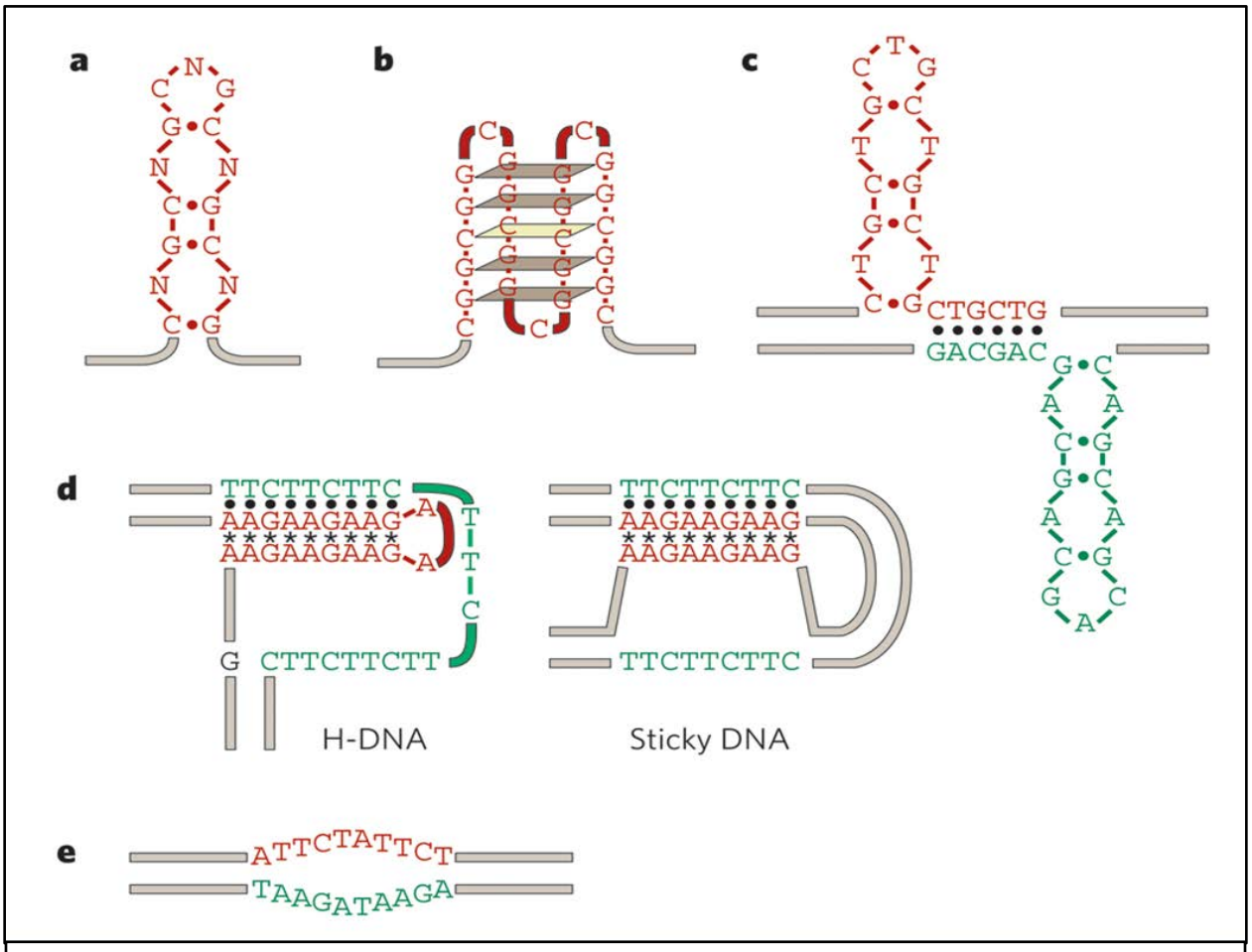



Figure 2. **Non-B DNA structures.** A. Single hairpin structure from CXG motif. B. G-quadraplex structure from GGC repeats. C. Model of slipped-strand DNA formed from CTG and CAG hairpins. D. Py-Pu triplex DNA forming H-DNA and sticky DNA. E. Easily unwound DNA (Adapted from Mirkin et al. (2007)).

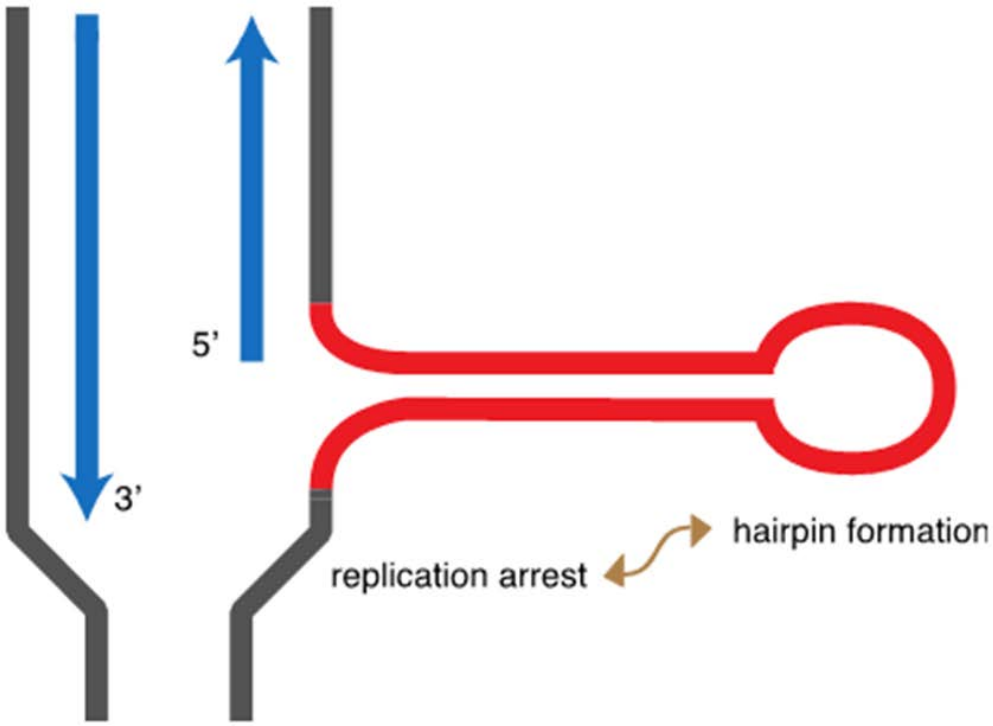



Figure 3. **Model showing hairpin formation leading to replication fork arrest.** This model demonstrates the formation of a hairpin structure which leads to the stalling of the replication fork.

**Mismatch Repair Proteins**

Various mechanisms have been proposed to contribute to the instability seen within microsatellites. This includes DNA replication, base excision repair (BER), double strand break (DSB) repair, nucleotide excision repair, recombination, mismatch repair (MMR), and replication slippage (Schmidt et al. 2016). Replication slippage involves slipped strand mispairing at a replication fork. It can account for the small scale contractions and expansions seen within microsatellites (Chen et al. 2005). Slipped-strand DNA can form during various metabolic processes including replication, repair, recombination, transcription and at unwound DNA (Axford et al. 2013).

The strongest candidate for expansion mechanism behind microsatellite expansions is likely the process of mismatch repair. While MMR usually functions as a means of preventing mutations, it is a likely culprit for causing mutations which lead to disease. MMR is unable to efficiently resolve hairpin structures that arise during replication, DNA repair, and transcription. This has been identified to be responsible for (CTG)•(CAG) expansions and has been shown in various model organisms wherein loss of MMR proteins blocks repeat expansion (Schmidt et al. 2016).

MSH2 is a tumor suppressor gene that encodes a mismatch repair protein. It forms a heterodimer with MSH6 to form the human MutSα mismatch repair complex. It also dimerizes with MSH3 to form the MutSβ DNA repair complex. These complexes are involved in various forms of DNA repair including transcription-coupled repair (Mellon I et al. 1996), homologous recombination, and base excision repair (Wind et al 1995,

Pitsikas et al 2007). MutSα is known to be more abundant in cells and important in recognizing single nucleotide misalignment, and MutSβ is needed to recognize larger insertion/deletion loops. MutSβ is important for processing hairpin structure formations (Gupta et al. 2012, Ellard 2017, SPIKE database 2011).

As such, the mismatch repair pathway is an important DNA repair pathway and plays a vital role in maintaining genomic stability. This pathway is also conserved in bacteria, yeast, and mammalian cells. Mutations within the MMR gene are associated with a variety of cancers, namely a predisposition to developing hereditary non-polyposis colorectal carcinoma (HNPCC) and Lynch syndrome (Pitsikas et al. 2007). In more recent studies it has been shown that the mismatch repair protein MSH2 levels are altered in non-melanoma skin cancers. MMRs interaction in DNA damage response to ultra violet stressors reveals a link between MMR and base excision repair (Pitsikas et al. 2007).

**Base Excision Repair Proteins**

DNA polymerase Beta (Polβ) is involved in base excision repair (BER) which is considered a gap-filling DNA synthesis. It is normally found in the cytoplasm but translocates to the nucleus when DNA damage occurs (NCBI POLB, Idriss et al. 2002). Base excision repair is responsible for deletion of damaged DNA bases and repairing of DNA single stranded breaks (Barakat et al. 2012). Reduction in the base excision repair pathway through Polβ deficiency has been linked to neurodegeneration,

haploinsufficiency, accelerated DNA damage and increased mutational response to carcinogens (Cabelof et al. 2003, Sykora et al. 2015).

**Crossover Junction Endonucleases**

Another mechanism involved in genome stability utilizes the crossover junction endonucleases. When replication forks stall at non-canonical DNA, the template needs to be processed to allow replication fork restart to occur (Ciccia 2003). Stalled or collapsed replications forks can be restarted through the process of homologous recombination repair (HRR). One of the proteins involved in this pathway is the MUS81 nuclease. MUS81 forms a complex with either EME1 or EME2 depending on what stage the cell cycle is in. MUS81-EME1 is active during G2/M phase, and MUS81-EME2 activates during S phase (Pepe and West 2013). MUS81-EME1 complex is required in the formation of double strand breaks after replication stalling, restarting of the replication process, cleavage of common fragile sites, and repair of interstrand cross-links (Xing 2015, Pepe 2014.) The MUS81-EME2 complex is thought to process non-S phase HRR D-loop structures by cutting the 3'-invading strand. MUS81-EME2 also acts on 5'-flap structures to cleave the duplex arm. It has also been found that in telomerase negative ALT (Alternative Lengthening of Telomeres) cells, MUS81-EME2 is needed for telomere maintenance (Pepe 2013).

While the heterodimers of MUS81-EME1 and MUS81-EME2 are both endonucleases, they have different cellular functions. MUS81-EME2 is not only

responsible for replication fork cleavage and restart in S-phase of the cell cycle; it is also able to cleave various DNA structures including Holliday junctions and nicked duplexes. MUS81-EME1 primarily functions in G2/M and is responsible for Holliday junction cleavage and fragile site expression (Amangyeld 2013, Pepe 2014). Figure 4 shows a summary of MUS81-EME1 and MUS81-EME2 heterodimers and their functions during cell cycle.

In a recent study, it was suggested that MUS81-EME1 and MUS81-EME2 have related but non-overlapping roles in DNA synthesis, with MUS81-EME2 having a more dominant and versatile role than MUS81-EME1. In this study, MUS81-EME2 was able to process intact Holliday junctions, nicked Holliday junctions, and nicked duplex and gapped duplex DNAs in order to form DSBs; while MUS81-EME1 only efficiently cleaved nicked Holliday junctions. In this study MUS81-EME2 showed higher activity in vitro for various types of DNA structures when compared to MUS81-EME1. MUS81-EME2 was shown to process a wider range of damaged non-B DNA structures that are difficult to repair by other downstream repair pathways (Amangyeld et al. 2014).

A study specifically looking at PARP inhibitors resistance in BRCA2-deficient ovarian cancers has also suggested that MUS81 deficiency enhances replication fork protection, leading to chemoresistance and poor cancer prognosis (Rondinelli et al. 2017). It is suggested that prolonged MUS81 activity at a stalled replication fork may be toxic. This can result in uncontrolled fork degradation or collapse (Rondinelli et al. 2017). The prolonged activity of MUS81 results in the restoration of replication fork progression, but

disables the double strand DNA break repair of homologous recombination. This restoration of fork protection is thought to lead to the chemoresistance seen in MUS81 deficient tumors (Schlacher 2017). MUS81 null cancer cells during unperturbed growth show a higher rate of DNA fork initiation, without activating novel replication origins, and slower replication fork progression (Fu et al. 2015).

MUS81 has been found to interact with various checkpoint proteins such as CHK1/CHK2, p53 and RAD51. It is also important in the survival of telomerase-negative cancer cells by maintaining the genetic integrity of telomeres through alternative lengthening of the telomeres pathway (Wu et al. 2010). Interestingly, while MUS81 deficient cells suffer increased replication stress sensitivity, MUS81 deficiency does not promote tumorigenesis; however, it is associated with poor cancer prognosis (Dendouga et al. 2005, Bartosova, Krejci, 2014). MUS81 downregulation has been connected significantly to hepatic metastasis in colorectal cancers. It has also been connected with poorer survival of patients compared with those with wild type MUS81 expression (Wu et al 2010).

It has been hypothesized that prolonged MUS81 activity at replication forks can lead to replication fork collapse (Rondinelli et al. 2017). This study also showed that decreased MUS81 activity at replication forks lead to fork stabilization. It has also been shown that MUS81 depletion reduces formation of double strand breaks, increases cell survival and allows S-phase progression (Forment et al. 2011). Another study shows that MUS81 has a role in the rate of replication fork progression. When MUS81 is absent,

DNA synthesis is slowed, suggesting a role of MUS81 in the pace of replication fork progression (Pepe et al. 2014).

In a study by Fugger et al. (2013) showed that MUS81 or EME1 depleted cells showed increased resistance to cytotoxic stress induced by hydroxyurea. Cells treated with siMUS81 and siEME1 exhibited impaired activation of the pro-apoptotic factor p53, and decreased double strand break formation following the treatment with hydroxyurea. In this study, it suggests that FBH1 helicase, which has been shown to respond to cellular replication stress, mediates activation of MUS81 while also regulating RAD51 activation to suppress homologous recombination. Their results showed that in cells treated with hydroxyurea, siRNA mediated knockdown of MUS81 or EME1 led to decreased double strand break formation. Their model suggests that MUS81 mediated DNA double strand breaks amplify checkpoint response proteins leading to apoptosis in order to eliminate cells with prolonged replication stress (Fugger et. al, 2013). These results also suggest that MUS81-EME1 may have a function during S phase.

In the present work, we describe a novel cell culture system in which we can model the in vivo effects of siRNA, and shRNA mediated knockdown of the various proteins mentioned previously, which are involved in microsatellite stability, replication fork progression and genome stability. This cellular model represents a diseased state expanded microsatellite which will be used as a model to show the effects of knocking down these proteins in the presence and absence of hydroxyurea induced replication stress. The following results indicate DNA repair pathways as a possible cause for

instability and double strand breaks seen within the DF2 Myc $CTG_{100}$ cells. Using this approach we show that expanded CTG microsatellites are also sites of replication-dependent chromosome breakage. Break-induced replication has been implicated in gross chromosomal rearrangements found in cancers. Thus, these results are relevant to understanding the basis for oncogenic genomic instability.

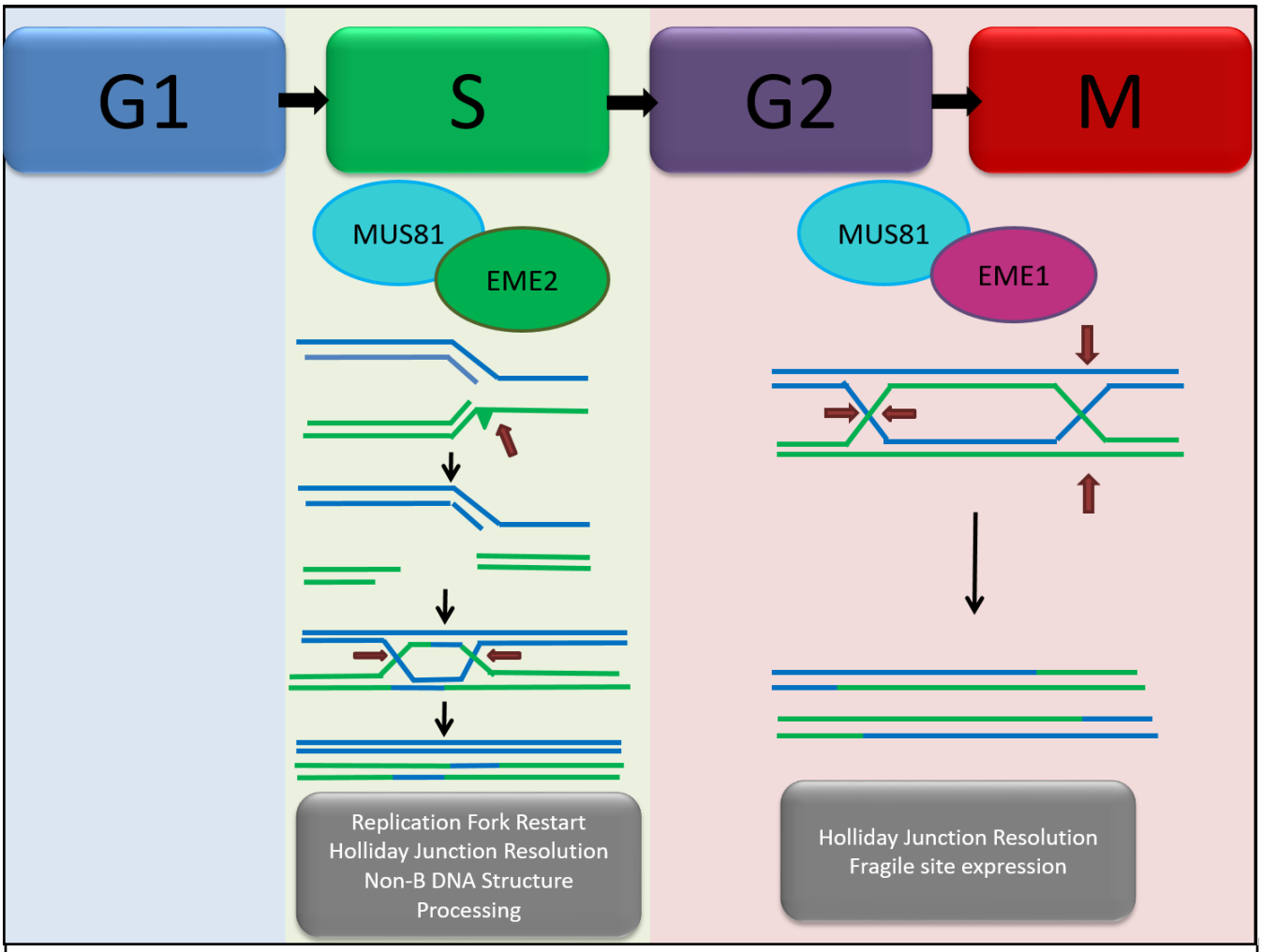



Figure 4. **Model for MUS81 EME1 and EME2 interaction.** MUS81 forms a complex with EME2 during S-phase of the cell cycle, this complex is responsible for replication fork restart and Telomere Maintenance. MUS81 forms a complex with EME1 during G2 and M phase of the cell cycle, this complex is responsible for Holliday junction resolution and fragile site expression. The arrows represent cleavage sites. The triangle represents a replication fork barrier.

# MATERIALS AND METHODS

## Cell Culture

Cells were derived from a HeLa based acceptor cell line (HeLa/406 cells) with a single FRT (FLP recombinase target) site (Figure 5). A plasmid containing FLP recombinase and a donor plasmid containing CTG/CAG repeats along with a c-Myc replicator were co-transfected into the HeLa/406 acceptor cells. This insertion site is located on chromosome 18, and is a single integration site. The repeats are placed adjacent to the c-Myc replicator so that the polarity of replication is known. The CTG repeats are placed on the template for lagging strand synthesis. The official name for this cell line is DF2 Myc $(CTG/CAG)_{100}$. This represents a diseased state microsatellite. The threshold frequency of trinucleotide repeats which can lead to a diseased state is over 50 repeats. 100 repeats of CTG/CAG were integrated into the cell line since this number falls well beyond threshold and can represent a diseased state microsatellite. Another cell line, DF2 Myc $(CTG/CAG)_{23}$ representing a healthy length microsatellite, was also made as a control. These cell lines were developed by Todd Lewis (Wright State University). These cell lines were grown and maintained in Dulbeco's Modified Eagle's Medium (DMEM) supplemented with 10% newborn calf serum and 5% $CO_2$.

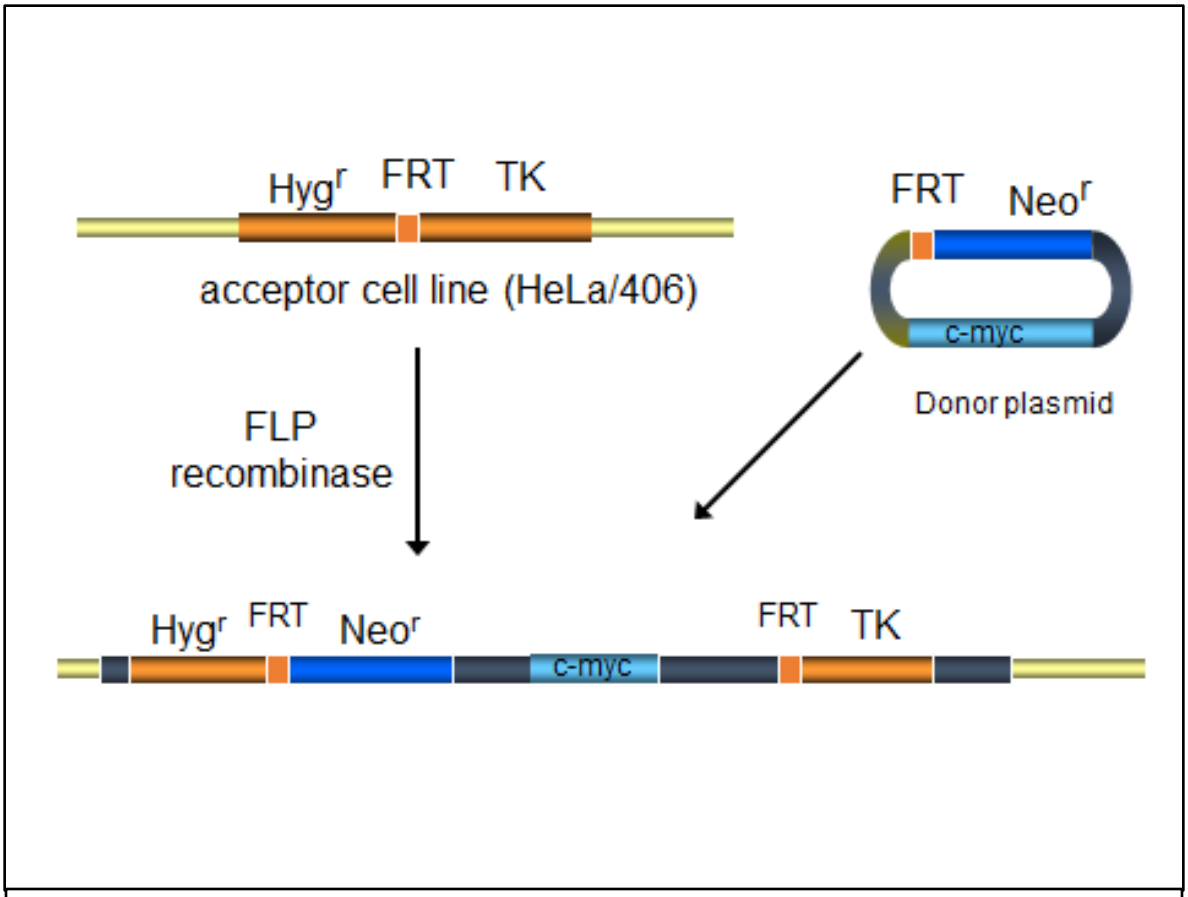



Figure 5. **Chromosomal integration of c-Myc origin constructs.** Diagram showing acceptor cell line with single FRT site and donor plasmid with c-Myc replicator/ the CTG/CAG repeats are right next to c-Myc replicator and form the basis of DF2 Myc (CTG/CAG) cell lines.

**DF2 Myc $(CTG/CAG)_{100}$**

The primary cell line used in the following experiments is the DF2 Myc $(CTG/CAG)_{100}$ cell line (also referred to as CTG100 cells). Dual fluorescence 2 (DF2) cell lines have two color marker genes, which are red (red fluorescent protein, dTomato) and green (Green fluorescence protein, eGFP), whose presence or absence can be analyzed by flow cytometry. The presence of both proteins will result in a yellow reading, or double positive, whereas a loss of dTomato will result in green readings. Likewise, loss of eGFP will result in red readings. The ectopic FRT integration site has three identical Alu sequences denoted Alu-1, Alu-2, and Alu-3. These are known sites for recombination. The c-Myc region follows Alu-2 and is an origin of replication. CTG/CAG repeats are placed adjacent to the c-Myc replicator. Tomato is placed between Alu-1 and Alu-2 while eGFP is placed between Alu-2 and Alu-3 (Figure 6). As mentioned previously, the trinucleotide microsatellite region is placed adjacent to the c-Myc region. These cell lines are designed so that if DNA undergoes breaks within the trinucleotide repeat region, it would lead to separation of the color markers genes. Depending on the type of damage and the repair mechanism and recombination patterns taken by the cell, we can observe variation within flow cytometry profiles.

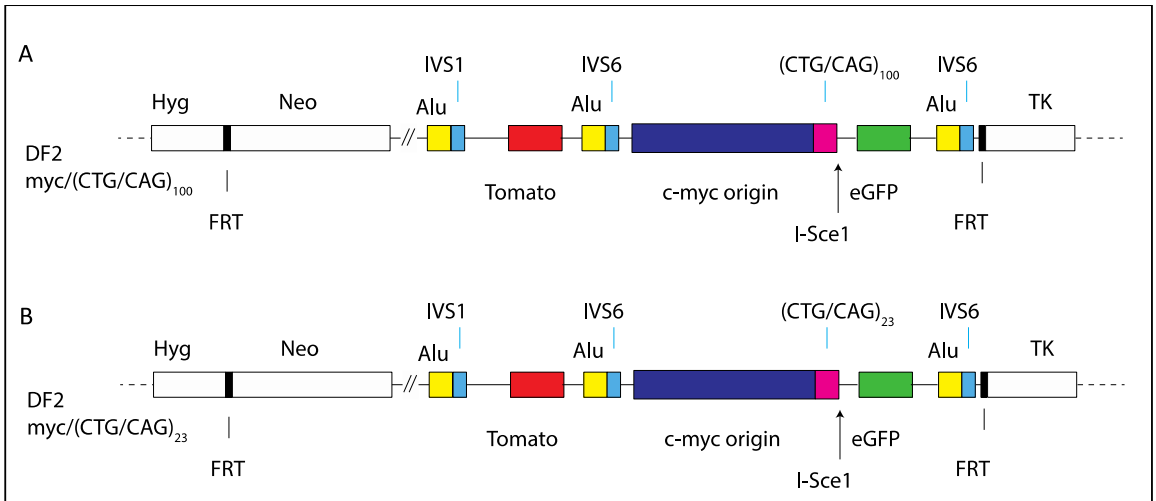



Figure 6. **Cell line constructs.** Map of DF2 Myc $(CTG/CAG)_{100}$(A), and DF2 Myc $(CTG/CAG)_{23}$ (B) cells. These cell lines will be referred to as CTG100 and CTG23 throughout this paper.

**siRNA treatment**

CTG100 cells, and CTG23 cells (Figure 6A and B) (40% confluent) were transfected with Lipofectamine 2000 Reagent (Thermo Fisher Scientific) and 100 nM (final concentration) of small interfering RNA (siRNA) targeting MUS81 (Dharmacon), EME1 (Invivogen), and EME2 (Invivogen), Polβ (Invivogen) in 6 well plates for either 2 (MUS81) tranfections 24 hrs and 96 hrs after initial plating of cells, or single (EME1, EME2, MSH2, MSH3, MSH6, Polβ) transfection 24 hours after initial plating of cells. Twenty-four hours after initial plating of cells. Forty-eight hours after initial plating of cells, $(CTG/CAG)_{100}$ were treated with 0.2 mM hydroxyurea (HU). Hydroxyurea is used to induce replication stress by inhibiting ribonucleotide reductase and decreasing dNTP levels. Control experiments were carried out using AllStars negative non-targeting siRNA (Qiagen).

Fully confluent 10 cm plates were trypsinized, suspended in 6 ml of DMEM and $4x10^5$ cells were put onto 6 well plates on day 0, allowing plates to be 30%-40% confluent for the initial siRNA knockdown transfection. Twenty-four hours after plating the cells onto a 6 well plate, wells were treated with siRNA (MUS81, EME1, EME2, or siRNA control). siRNA transfection was performed as follows. In a 1.5 ml Eppendorf tube 250 ul OPTIMEM and 10 ul Lipofectamine2000 was mixed and incubated for 5 minutes. In a second Eppendorf, 250 ul of OPTIMEM was mixed with 10 ul of the respective siRNA (20 uM). After 5 minutes, the Lipofectamine2000 mix was added to the

siRNA mix and incubated for 20 minutes. The 6 well plates had 1.5 ml of 20% FBS DMEM media added to them, without antibiotics. After 20 minutes incubation the siRNA mix was added to the cells and incubated at 37 $^{o}$C. The final concentration of the siRNA was 100 nM. Twenty-four hours after siRNA transfection, cells were treated with 0.2 mM hydroxyurea for 4 days after which they were allowed to recover for 4 days and then were analyzed by flow cytometry and/or collected for DNA extraction. The 4 days of recovery is important because it allows the preexisting red and green proteins to degrade. This allows for the proper separation of colors by flow cytometry.

**shRNA treatment**

CTG100 cells (40% confluent) were transfected with Lipofectamine 2000 Reagent (Thermo Fisher Scientific) and 4 ug/mL of short hairpin RNA (shRNA) plasmid targeting MSH2, MSH3, and MSH6 in 6 well plates for single transfection 24 hours after initial plating of cells. Forty-eight hours after initial plating of cells, $(CTG/CAG)_{100}$ cells were treated with 0.2mM hydroxyurea (HU). Control experiments were carried out using AllStars negative non-targeting siRNA (Qiagen).

Fully confluent 10 cm plates were trypsinized, suspended in 6 ml of DMEM and $4x10^5$ cells were put onto 6 well plates on day 0, allowing plates to be 30%-40% confluent for the initial shRNA knockdown transfection. Twenty-four hours after plating the cells onto a 6 well plate, wells were treated with shRNA (MSH2, MSH3, MSH6). shRNA transfection was performed as follows. In a 1.5 ml Eppendorf tube 250 ul

OPTIMEM and 10 ul Lipofectamine2000 was mixed and incubated for 5 minutes. In a second Eppendorf, 250 ul of OPTIMEM was mixed with shRNA plasmid. After 5 minutes the Lipofectamine2000 mix was added to the shRNA mix and incubated for 20 minutes. The 6 well plates had 1.5 ml of 20% FBS DMEM media added to them, without antibiotics. After 20 minutes incubation the shRNA mix was added. The final concentration of shRNA was 4ug/mL. Twenty-four hours after shRNA transfection cells were treated with 0.2 mM hydroxyurea for 4 days after which they were allowed to recover for 4 days and then were analyzed by flow cytometry and/or collected for DNA extraction.

**PCR**

Sensitive detection and quantification of the frequency of mutant microsatellite sequences can be seen by using small pool PCR. Small pool PCR is an end-point PCR method which allows for an increased estimate of a low abundance template. This has advantages over a higher abundant template that has plateaued amplification early in the PCR. This PCR technique is then used on a number of replicates, to increase the chances of seeing mutations in the collected samples Using spPCR analysis of genomic DNA was performed with an internal control primer set to assess differences between cells in the population, as well as detect expansions, contractions or double strand breaks of the microsatellite sequences of interest.

Small pool PCR (spPCR) was performed using 50 pg-100 pg of genomic DNA (around 1-10 genome equivalents). Three to four small pool PCR reactions were performed on DNA template from each experimental condition. All PCR reactions occurred in the presence of HotStar Taq polymerase Master Mix (Qiagen) with every reaction. Amplification conditions for the ectopic site were 95 °C for 15 minutes, 94 °C for 1 minute, 55 °C for 30 seconds, 72 °C for 45 seconds, and 72 °C for 5 minutes. Steps two through four were repeated for forty-five cycles. Amplification conditions used for endogenous DMPK were 95 °C for 15 minutes 94 °C for 1 minute, 58 °C for 1 minute, 72 °C for 1 minute and 72 °C for 10 minutes; steps two though four were repeated for 35 cycles. PCR products were resolved on an 8% polyacrylamide gel, and stained with GelRed (Biotium). Images were obtained on a Fuji LAS-3000. Primers used are listed in Table 1.

| Primer | Sequence | Product Length |
|---|---|---|
| Blue-F | AAAAAAGATCCTCTCTCGCTAATCTCC | 515 bp/ 285 bp |
| SFFV-R | GAACTTCCCTATTCTTGGTTTGG | |
| 406-R | GGAGGATGGAACACGGACGG | 125 bp |
| 409-F | GAAGGGTCCTTGTAGCCGGGAA | |
| DMPKUp-F | CTCACTGGTCACTGGTTTCTT | 325 bp |
| DMPKUp-R | CCAACCCAACTTCATCCTCTAC | |
| Tom-F | CGAGGAGGTCATCAAAGAGTTC | 812 bp |
| BGH-R | TAGAAGGCACAGTCGAGG | |
| GFP-F | ACGTAAACGGCCACAAGTTC | 641 bp |
| GFP-R | GTCCATGCCGAGAGTGATC | |
| Lac-F | CTTCAAATCCGACCCGTAGA | 8.1 kb |
| TK-R | GTAAGTCATCGGCTCGGGTA | |
| 8937951 | GTAGGTTCAAAGGGTGGGT | 97 bp |
| 8937698 | GGCAACGTGACAAGGAATG | |

Table 1. **Primers used for spPCR and PCR**

**siRNA treatment and Western blot**

The protocol mentioned above for siRNA treatment was followed with the respective cell lines and siRNAs. Cells were harvested at 24 hours, and 48 hours after transfection to check for protein knockdown by western blot. Whole cell lysates from treated or untreated $CTG_{23}$ cells were prepared. After sodium dodecyl sulphate-polyacrylamide gel electrophoresis, membranes were probed for the respective proteins to check that the treatment was successful. EME1 antibody (Abcam: ab88878) was used, at a 1:500 dilution. MUS81 antibody (Abcam: ab14387) was used at a 1:500 dilution. EME2 antibody (Abcam: ab129443) was probed at a 1:100 dilution. γH2AX (Millipore: 05-636) was used at a 1:20,000 dilution to check for change in baseline cellular DSB formation. β-actin antibody (Sigma-Aldrich A5441) used at a 1:1000 dilution was used as a loading control. After incubation with secondary antibody membranes were imaged on a Fuji LAS-3000 using ImageReader.

**Western blot and Hydroxyurea Sensitivity**

HeLa cells were treated either with 2 mM HU for 6 hours (positive control), not treated (negative control), or treated with 0.2 mM HU and collected every other day for 10 days. Whole cell lysates were prepared. After SDS gel electrophoresis membranes were probed with various antibodies to check for DNA damage response. pCHK1 antibody (Cell Signaling) was used at a 1:500 dilution. pCHK2 antibody (Cell Signaling) was used at a 1:500 dilution. γH2AX antibody (Millipore: 05-636) was used at a 1:20,000

dilution. β-actin antibody (Sigma-Aldrich A5441) used at a 1:1000 dilution was used as a loading control.

**Immunofluorescence and Hydroxyurea Sensitivity**

HeLa cells were treated with HU as mentioned previously. Cells were affixed to a glass slide using 2% para-formaldehyde, and 0.2% of Triton X solution was used to improve antibody permeability. Primary antibody of 1:5000 solution γH2AX (Abcam: ab11174) was used. DAPI (1:25000 dilution) was used to visualize cell nuclei. Invitrogen A11036 Alexa Fluor 568 was used as a secondary antibody. Images were taken on the FV1000 confocal microscope.

**Flow Cytometry**

After cells were treated with siRNA and hydroxyurea to induce replication stress, they were allowed to recover in DMEM medium with 10% NCS and antibiotics. Cells were then harvested to analyze their flow cytometry profiles. Harvested cells were aliquoted in 1.5 ml Eppendorf tubes and centrifuged (Eppendorf Centrifuge 5415D) at 300 rpm for 3 minutes. Media was aspirated and cells were washed with 300 ul cold PBS, and were re-centrifuged at 300 rpm for 3 minutes. After a final wash of 200 ul cold PBS, cells were analyzed using a C-Flow Plus Accuri Cytometer ®. Figure 7 shows how the results will be represented by the flow cytometer; showing readings from the DF2 Myc CTG100 and the DF2 Myc CTG23 cell lines. Flow cytometry readings included 20,000

events, with gating placed around the densest number of cells. The top right quadrant contains the double positive cells (yellow), top left the dTomato cells (red), bottom right showing the eGFP cells (green) and bottom left showing the double negative cells which have lost both color markers. In the readings acquired using the flow cytometer a difference of 5% or more is considered to be significant.

**Flow Cytometry to Detect Cell Cycle Phase**

After siRNA treatment and collection of samples after 24 and 48 hours, cells were harvested in 1.5 ml Eppendorf tubes and centrifuged at 500 rpm for 3 minutes. Media was aspirated off and washed with cold PBS, then spun down at 500 rpm for 3 minutes. Cells were then treated with 70% ethanol and permeabilized at -20 $^{o}$C overnight. Cells were then centrifuged at 1500 rpm for 5 minutes and the ethanol was aspirated. This was followed by washing the cells with 1 ml cold PBS and adding 0.75 ul of 100mg/ml RNAse and incubating at 37$^{o}$C for 20 minutes. Lastly DRAQ7 (Cell Signaling) dye was added at a final concentration of 7.5 uM and incubated at room temperature in the dark for 30 minutes, after which cells were analyzed using C-Flow Plus Accuri Cytometer. DRAQ7 dye was used in lieu of propidium iodide in order to reduce false reads from dTomato red fluorescence protein.

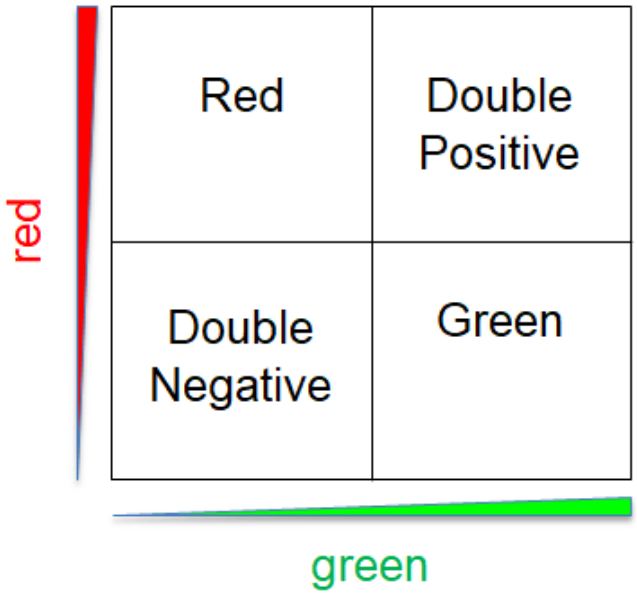



Figure 7. **Flow Cytometry chart.** Results produced by C-Flow Plus Accuri Cytometer will be represented in this fashion. All results are with the DF2 Myc CTG100 or the DF2 Myc CTG23 cell lines.

## Results

### I. Hydroxyurea is effective agent as replication stress and DSB inducer

It has been shown in previous studies that the DF2 Myc $(CTG)_{100}$ cells are sensitive to replication stressing agents. Both Figure 8 and Figure 9 were used to show the efficiency of using hydroxyurea as a replication stressing agent. HeLa cells were used in both of these experiments. In Figure 8, 2 mM HU was used for 6 hours to produce a positive control, and negative control was non-treated cells. Other cells were treated with 0.2 mM HU and collected after 2, 4, 6, 8, and 10 days of treatment. γH2AX antibody was used as an indicator of DSB and replication stress within the cells, DAPI was used as a control to visualize cell nuclei. Figure 8 immunofluorescence shows increase in γH2AX signal as the number of days increase with treatment. In Figure 9 HeLa cells were harvested at the same intervals with the same treatment. Protein was extracted from these cells to corroborate the findings from the immunofluorescence. The western blot allowed for looking at the treated cells from a population stand point as well as using antibodies to check for other signs of DNA damage response signaling. pCHK1, pCHK2 and γH2AX were used to show the increase in DNA damage response signaling as the treatment with HU continued through the time test. This shows that low dose hydroxyurea treatment is an effective drug to use to elicit replication stress.

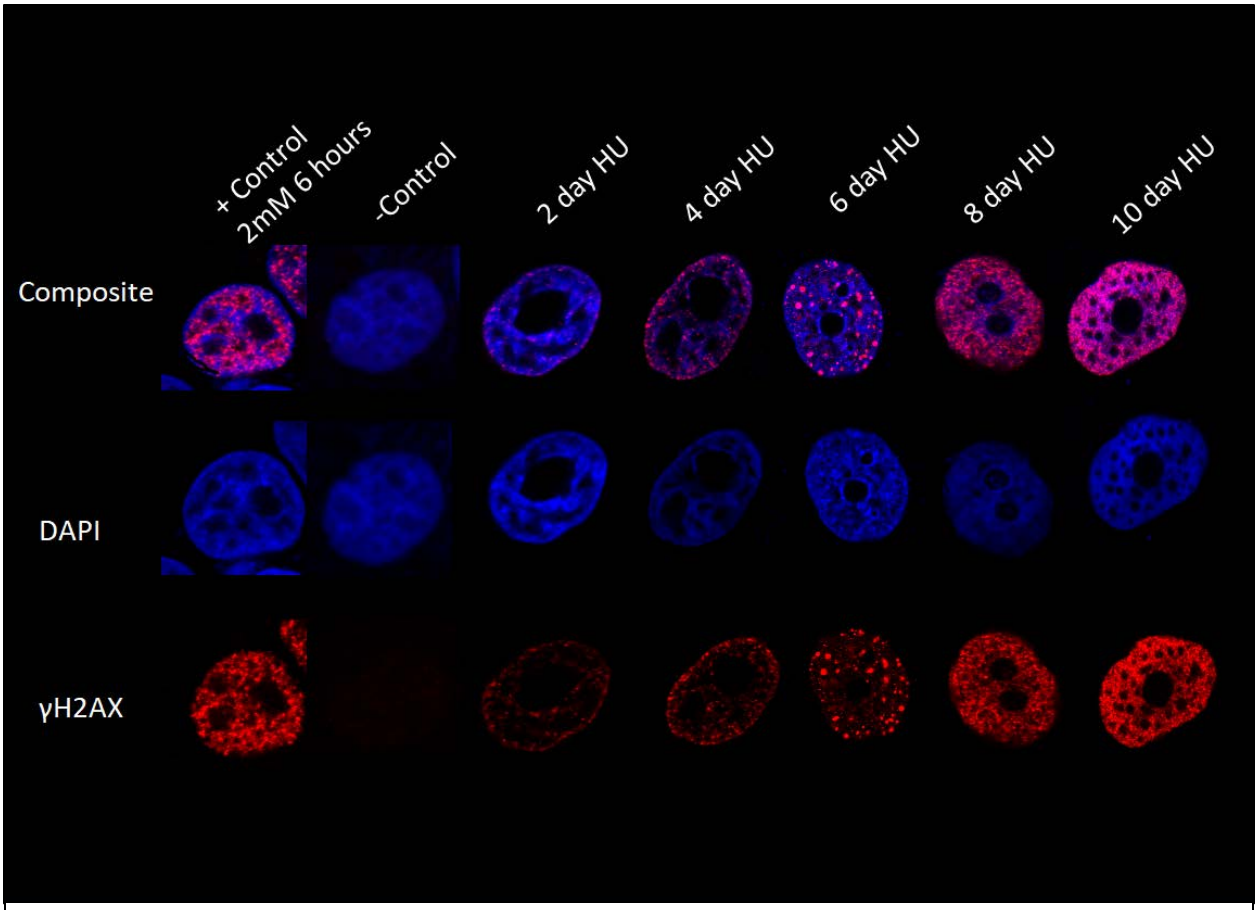



Figure 8. **Hydroxyurea is effective agent for inducing DNA replication stress.** Immunofluorescence of HeLa cells treated with hydroxyurea to induce replication stress and DSB. Positive control was treated with 2 mM HU for 6 hours. Cells were collected in two day intervals after being treated with 0.2 mM. Antibody against γH2AX was used as indicator of replication stress with DAPI as a control.

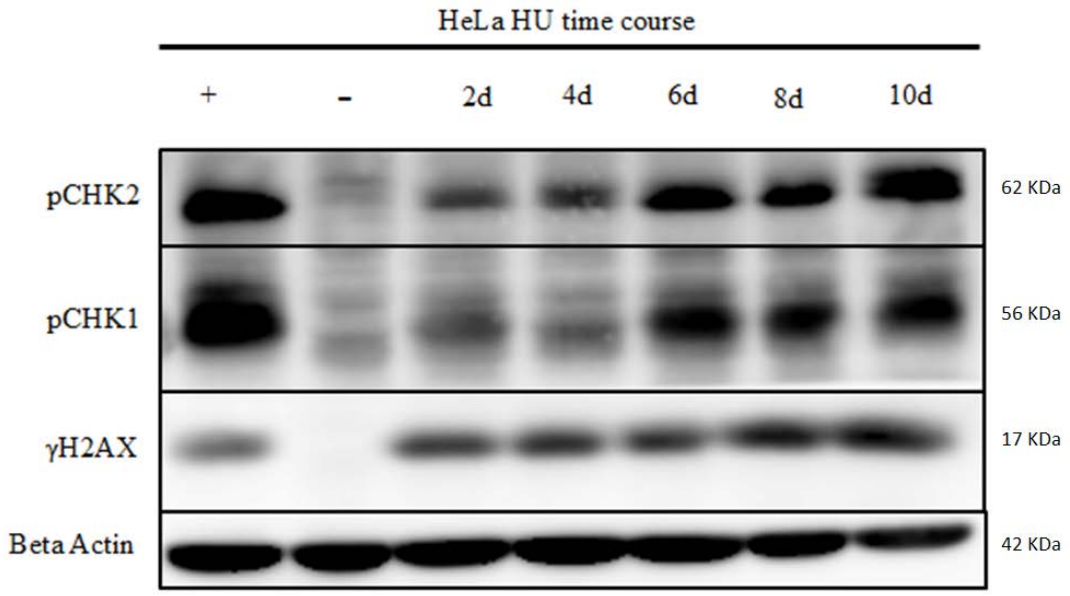



Figure 9. **Western blot showing DNA damage response in presence of hydroxyurea.** HeLa cells were treated as mentioned previously, protein was harvested at the time intervals and probed for pCHK2, pCHK1, and γH2AX, increasing in signal as the time course experiment progressed. Beta actin was used as a loading control.

## II. Knockdown of Mismatch Repair Proteins in DF2 Myc $CTG_{100}$ cells affect DNA damage tolerance

It was mentioned previously that MSH2 is a DNA mismatch repair protein; it is a tumor suppressor and a caretaker gene. It dimerizes with MSH6 to form the MutSα complex or dimerizes with MSH3 to form MutSβ. These proteins are important for maintaining the stability of microsatellites by base mismatch repair of short insertion/deletion loops. In previous studies with the CTG100 cell lines, it has been shown that these cells are sensitive to replication stress induced by hydroxyurea and are unstable. This lead to the interest in exploring if the CTG100 cells overcome replication stress by means of mismatch repair proteins. This assay was performed with shRNA to knock down MSH2, MSH3, and MSH6. Along with these various knockdowns, we combined the knockdown with replication stress using hydroxyurea. The controls for this assay were siControl as well as a siControl with hydroxyurea treatment. This way each treatment was accompanied by a control. The use of siControl essentially meant there was no protein knockdown. Results for these assays are shown in Figures 10-12.

Figure 10 shows flow cytometry profiles representing results for knockdown and no knockdown of MSH2 in presence or absence of hydroxyurea. In shMSH2 treated cells double positive (yellow) cells counted at 58.6% while control (untransfected) yellow cells were 52.4%. GFP cells for shMSH2 treated cells were 29.3% while control GFP cells were 34.6%. When cells were treated with hydroxyurea, control double positive cells

decreased to 41.1% and GFP cells increased to 48.4%, while shMSH2 treated cells with hydroxyurea decreased double positive cells to 53.7% and increased GFP cells to 37.3%. An increase in GFP signal (increase in green cells) from the flow cytometer indicates an increase in double strand breaks. A decrease in GFP signal indicates a decrease in double strand breaks. Not only did the treatment with shMSH2 increase the overall double positive cells and decrease the GFP cell number, the overall increase of the GFP was reduced from the control. These results show that the CTG100 cells tolerance for DNA damage and replication stress is increased when treated with shMSH2. Knockdown of MSH2 reduces the percentage of green cells (decreases the amount of double strand breaks) in the absence or presence of hydroxyurea, suggesting that MSH2 knockdown stabilizes the ectopic CTG microsatellite. The conclusion that the decrease in MSH2 protects the CTG microsatellite from DSBs is consistent with a previous study by Guo et al. (2017) which shows that MSH2/3 (MustSβ) contributes to microsatellite instability.

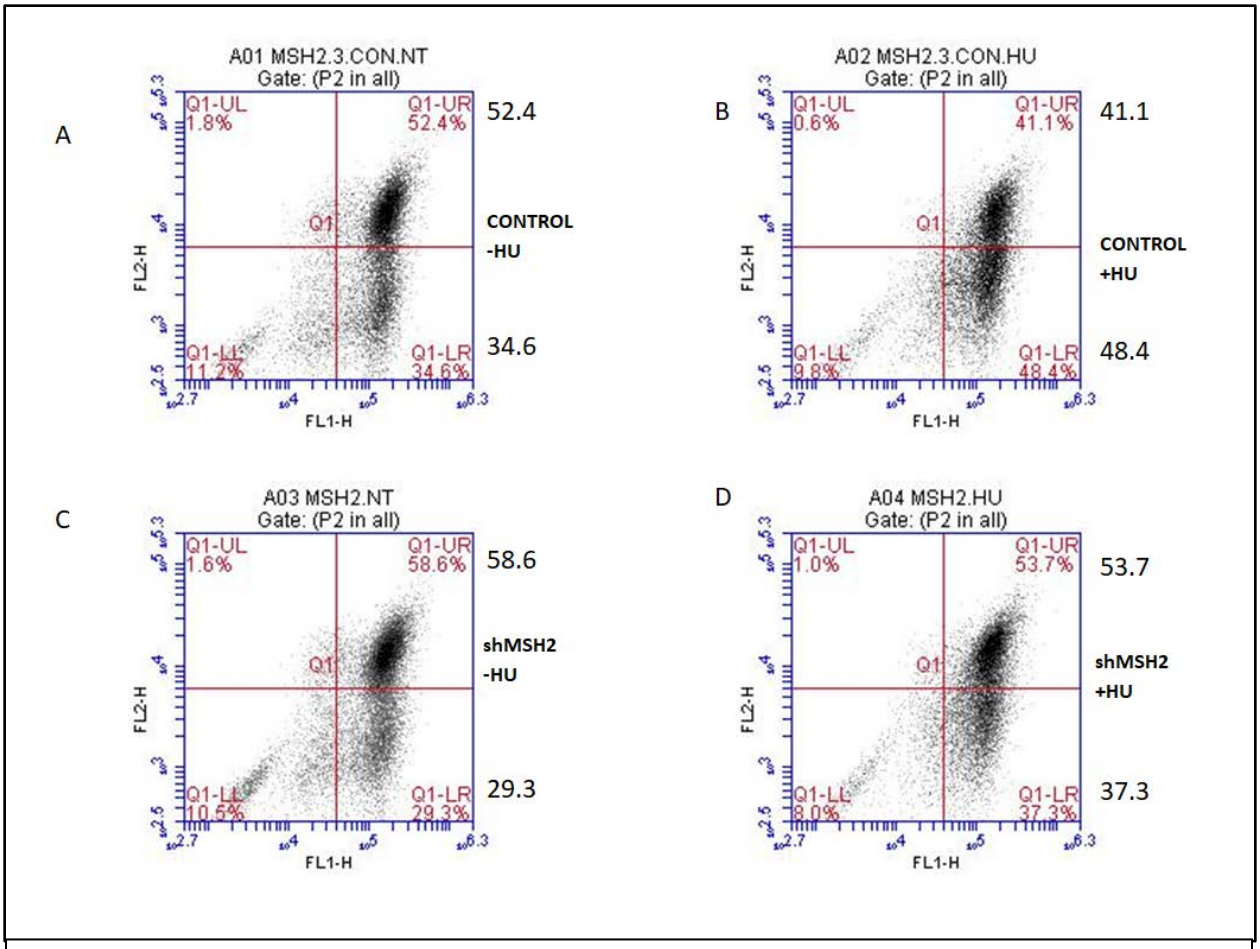



Figure 10. **Flow cytometry of cells treated with shMSH2**. (A) Control CTG100 cells not treated with HU. (B) Control CTG100 cells treated with 0.2mM HU. (C) CTG100 cells treated with shMSH2 without HU. (D) CTG100 cells treated with shMSH2 and HU. Treatment with hydroxyurea causes an increase in green cells indicating occurrence of double strand breaks. However, the knockdown of MSH2 reduces the number of green cells present when treated with HU.

MSH2 forms the MutSβ complex with MSH3. To test the prediction that the MutSβ is involved in destabilization of the CTG microsatellite, DF2 Myc $CTG_{100}$ cells were transfected with shMSH3 (Figure 11). Control (untransfected cells) cells without hydroxyurea treatment show 52.4% yellow cells and 34.6% green cells. With untransfected controls, cells treated with hydroxyurea shows a decrease of yellow cells to 41.1% and increase of green cells to 48.4%, indicating an increase in double strand breaks due to HU. CTG100 cells treated with shMSH3 show an increase in yellow cells to 63.0% and decrease of green cells to 25.3%, and when treated with both shMSH3 and hydroxyurea, double positive cells decrease to 54.1% and green cells increase to 35.7%. The effects of shMSH3 knockdown are similar to the protective effects of shMSH2, in either the presence or absence of hydroxyurea.

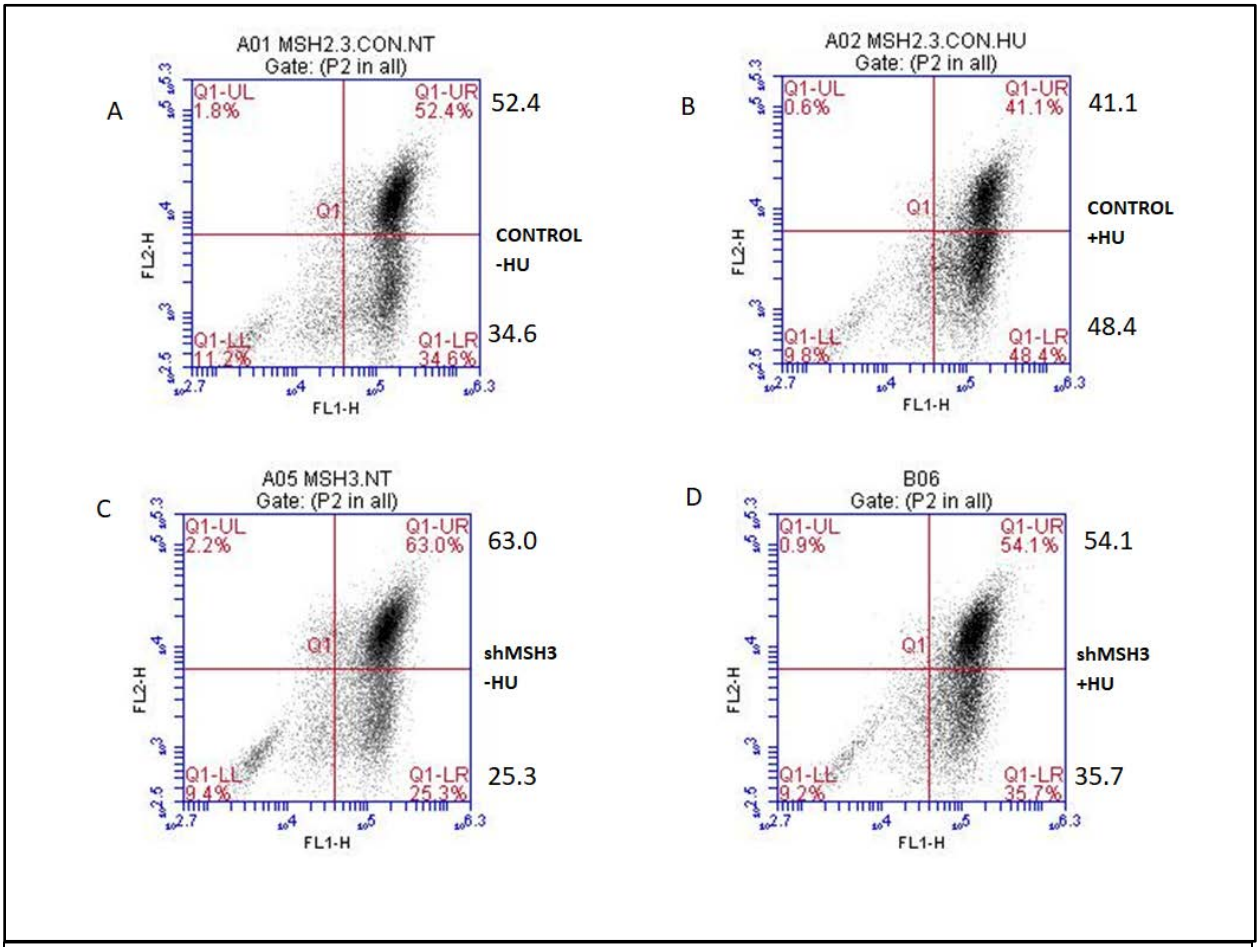



Figure 11. **Flow cytometry of cells treated with shMSH3. (**A) Control CTG100 cells not treated with HU. (B) Control CTG100 cells treated with HU. (C) CTG100 cells transfected with shMSH3 with no HU. (D) CTG100 cells transfected with shMSH3 and treated with HU. shMSH3 decreases green cells, +/- HU indicating a decrease in CTG100 double strand breaks.

MSH2 also forms the MutSα complex with MSH6, however $MSH6^{-/-}$ mice do not show increased or decreased CTG/CAG microsatellite instability (Foiry et al. 2006). Since MutSα has not been implicated in microsatellite instability, it is predicted that knockdown of MSH6 would have little effect in the present flow cytometry assay. Figure 12 shows the flow cytometry results from DF2 Myc $CTG_{100}$ cells treated with shRNA against MSH6 in presence or absence of hydroxyurea. In Figure 12, control cells show 47.2% cells double positive, and 38.8% of the cells green. When treated with hydroxyurea, control yellow cells fall to 38.5% and GFP cells increase to 51.4%. When treated with shMSH6 without hydroxyurea, double positive cells increase to 54.1% and GFP cells decrease to 33.0% from control. When treated with shMSH6 and hydroxyurea, double positive cells decrease slightly to 52.6% and GFP cells increase slightly to 36.2%. Thus, contrary to expectation, treatment with shMSH6 stabilized the CTG100 microsatellite against endogenous replication stress, and reduced the induction of CTG double strand breaks by hydroxyurea.

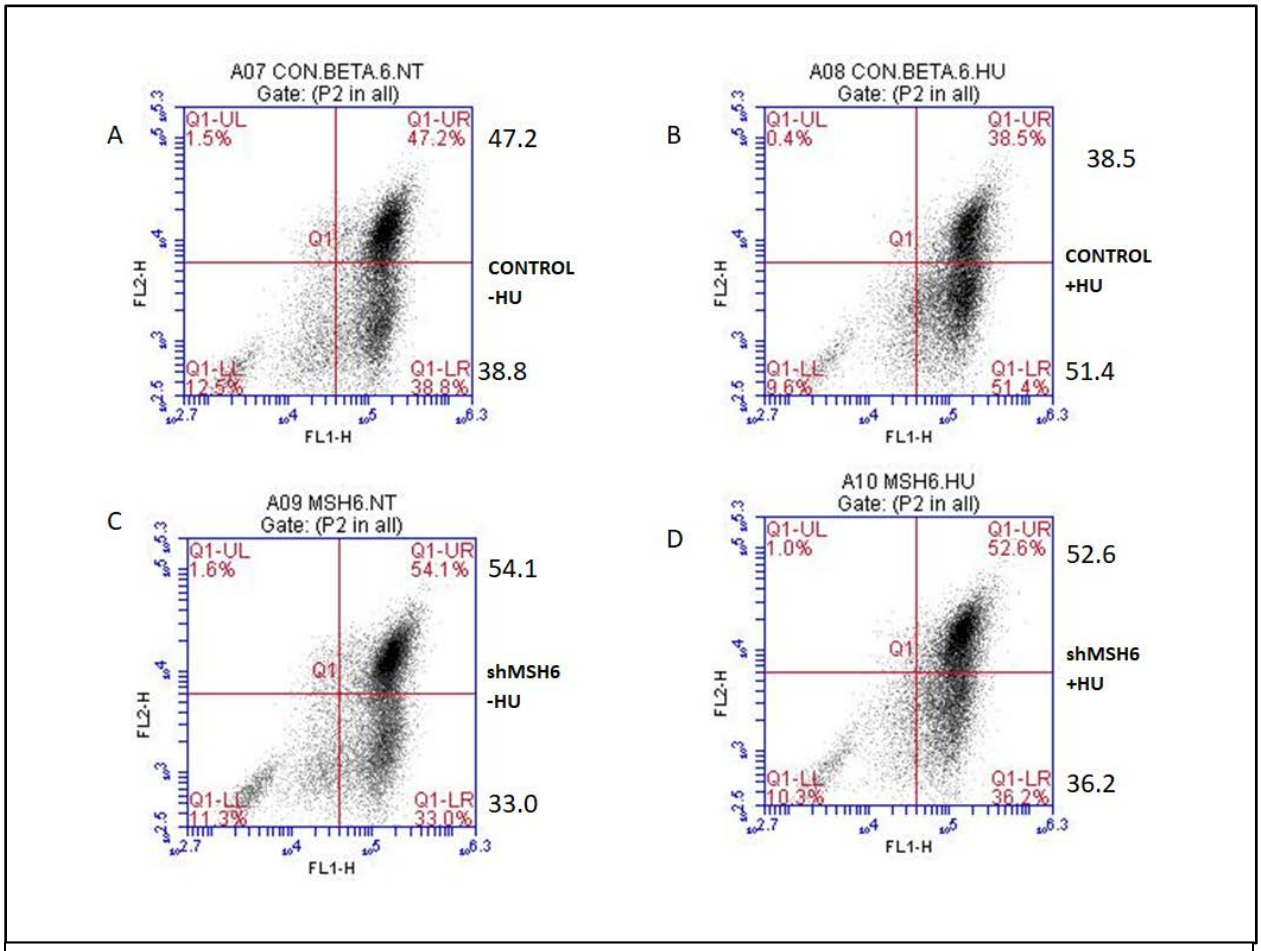



Figure 12. **Flow cytometry of cells treated with shMSH6.** (A) Control CTG100 control cells not treated with HU. (B) Control cells treated with HU. (C) CTG100 cells transfected with shMSH6 without HU. (D) CTG100 cells transfected with shMSH6 and treated with HU. MSH6 knockdown protects against hydroxyurea double strand breaks. When combined MSH6 and hydroxyurea treatment the HU effect diminishes.

## III. Loss of base excision repair protein Polβ affects the DNA replication stress response.

As mentioned previously, DNA polymerase beta (Polβ) is responsible for performing base excision repair (BER) which is required for DNA maintenance, replication, and recombination. Because Polβ has been shown to be recruited by MutSβ and act downstream of MSH2, we thought it was important to look at its effects on the DF2 Myc $CTG_{100}$ cell line when knocked down with siRNA. As before with the MSH2, MSH3, and MSH6 assays, CTG100 cells were treated with siPolβ with or without hydroxyurea. This was accompanied with siControl and siControl with hydroxyurea treatment. Results for this assay are shown in Figure 13.

The flow cytometry profile from Figure 13 shows data for knockdown or no knockdown of Polβ in presence or absence of hydroxyurea. In the control, 78.1% of cells are found to be double positive and 11.0% are GFP positive. When treated with hydroxyurea, control cell (untransfected) double positive numbers decrease to 60.6% and GFP positive cells increase to 23.3%. When treated with siPolβ, double positive numbers increase slightly to 77.4%, the GFP positive cells staying about the same at 11.7%. However, when treated with siPolβ and hydroxyurea, double positive cells increased to 72.2% and GFP cells have a dramatic decrease to 16.5%, compared with control. Therefore, treatment with siPolβ did not affect spontaneous CTG instability, but

protected against hydroxyurea induces double strand breaks. These results suggest that knockdown of the base excision repair pathway has little effect on endogenous levels of replication stress, but could contribute to DSBs at higher levels of replication stress in the expanded microsatellite. This is consistent with the role of MMR (MutSβ) and Polβ in microsatellite instabiltiy (Guo et al., 2017).

Although the results obtained with the shRNA treatment of MSH2, MSH3, MSH6, and Polβ exceeded the 4-5% threshhold of biological significance that we have found with this system, these are preliminary results that will need to be repeated to attain statistical significance. At this point, the data presented show a pattern but will need to be further explored to show definitive results.

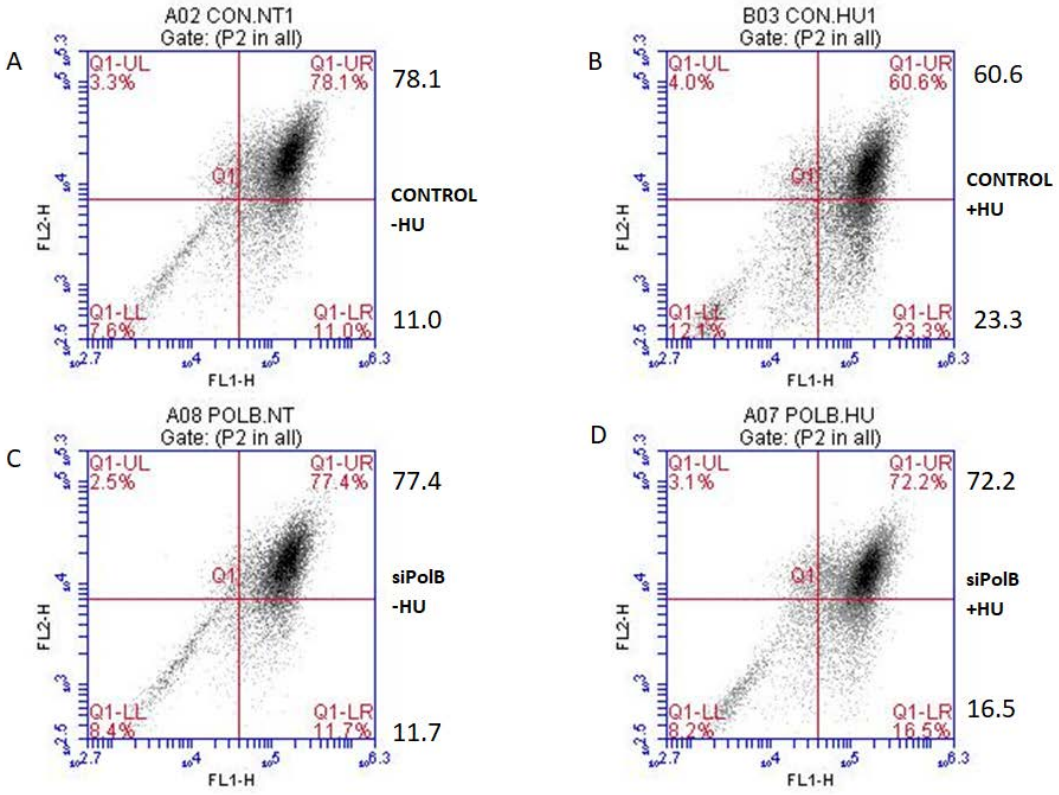



Figure 13. **Flow cytometry of cells treated with siPolβ.** (A) Control CTG100 cells without HU treatment. (B) Control CTG100 cells treated with HU. (C) CTG100 cells treated with siPolβ without HU treatment. (D) CTG100 cells treated with siPolβ and with HU. Polβ siRNA reduces the effect of hydroxyurea.

## IV. Loss of crossover junction endonuclease affects DNA damage tolerance

As mentioned earlier, MUS81 is a crossover junction endonuclease and is involved in the resolution of reversed replication forks, and also plays an essential role in completion of homologous recombination. MUS81 forms a complex with either EME1 or EME2 to process D-loops and Holliday junctions. Because resolving Holliday junctions is important in recombination and double strand break repair, we looked at how knocking down MUS81, EME1, and EME2 would affect the DF2 Myc $CTG_{100}$ cells. These knockdowns were also accompanied by inducing replication stress with hydroxyurea. This was combined with using siControl along with siControl and hydroxyurea treatment. Results for these assays are shown in Figures 14-17.

Figure 14 shows the flow cytometry profiles of CTG100 cells with or without knockdown of MUS81, in the presence or absence of hydroxyurea. Control cells were double positive at 75.6% and 18% are found to be eGFP positive. Under the replication stress caused by hydroxyurea, the control cells double positive decrease to 43.0%, and eGFP increases to 50.3%. When treated with siMUS81, the double positive cells increase to 82.7% and eGFP signal falls to 13%. When the knockdown of MUS81 was combined with hydroxyurea, the double positive cells decrease to 69.8% and the eGFP cells increase to 25.7%. The increase of eGFP cells in the siControl was 32.3%. In the siMUS81 treated cells, the eGFP count increased by only 12%. siMUS81 protects against spontaneous double strand breaks, and hydroxyurea induced double strand breaks (Quinet et al. 2017).

Figure 15 shows the same assay as Figure 14; the only difference in this assay was that we used DF2 Myc $CTG_{100}$ cells that were undergoing spontaneous DNA double strand breaks and did not have as high of a percentage of double positive cells at the beginning of the experiment. The double positive cells for the control are 52.9% and the eGFP positive cells were at 42.7%. When treated with hydroxyurea, the control cells double positive number decreased to 45.0% and the eGFP positive cells increased to 50.6%. However, when these cells were treated with siMUS81, the double positive cells drastically increased to 84.8% and the eGFP cells decreased to 13.2%. The combination of siMUS81, along with hydroxyurea, decreased the double positive cells to 73.0% and the eGFP cells increased to 23.6%. The overall increase in eGFP cells after treatment with hydroxyurea in this rescue experiment was very similar. The control was 7.9% increase in eGFP, and in the siMUS81 cells the eGFP cells increased by 10.4%. siMUS81 significantly protects against double strand breaks in the absence or presence of hydroxyurea.

The results from both Figure 14 and Figure 15 seem to indicate that under normal replication conditions, knocking down MUS81 has a pronounced stabilizing effect on the CTG repeats in these cells. These results imply that in the absence of MUS81 nuclease activity, there are fewer DSBs introduced at reversed forks near the CTG microsatellite. The knockdown of MUS81 in healthy cells which are undergoing fewer double strand breaks, seems to offer a stabilizing effect when treated with HU as well. There are significantly fewer cells shifting to green (fewer double strand breaks occurring) when

these healthier cells are treated with HU (Figure 14). However, when unhealthy cells which are undergoing a higher rate of double strand breaks are treated with HU induced replication stress, its stabilizing effect is minimized because the overall increase in eGFP cells is similar, despite the difference in double positive cells (Figure 15). Knocking down the MUS81 protein in the CTG100 cells essentially halted the spontaneous double strand breaks that were occurring within this cell line. These results suggest a toxic effect of MUS81 in cells with expanded microsatellites during normal replication.

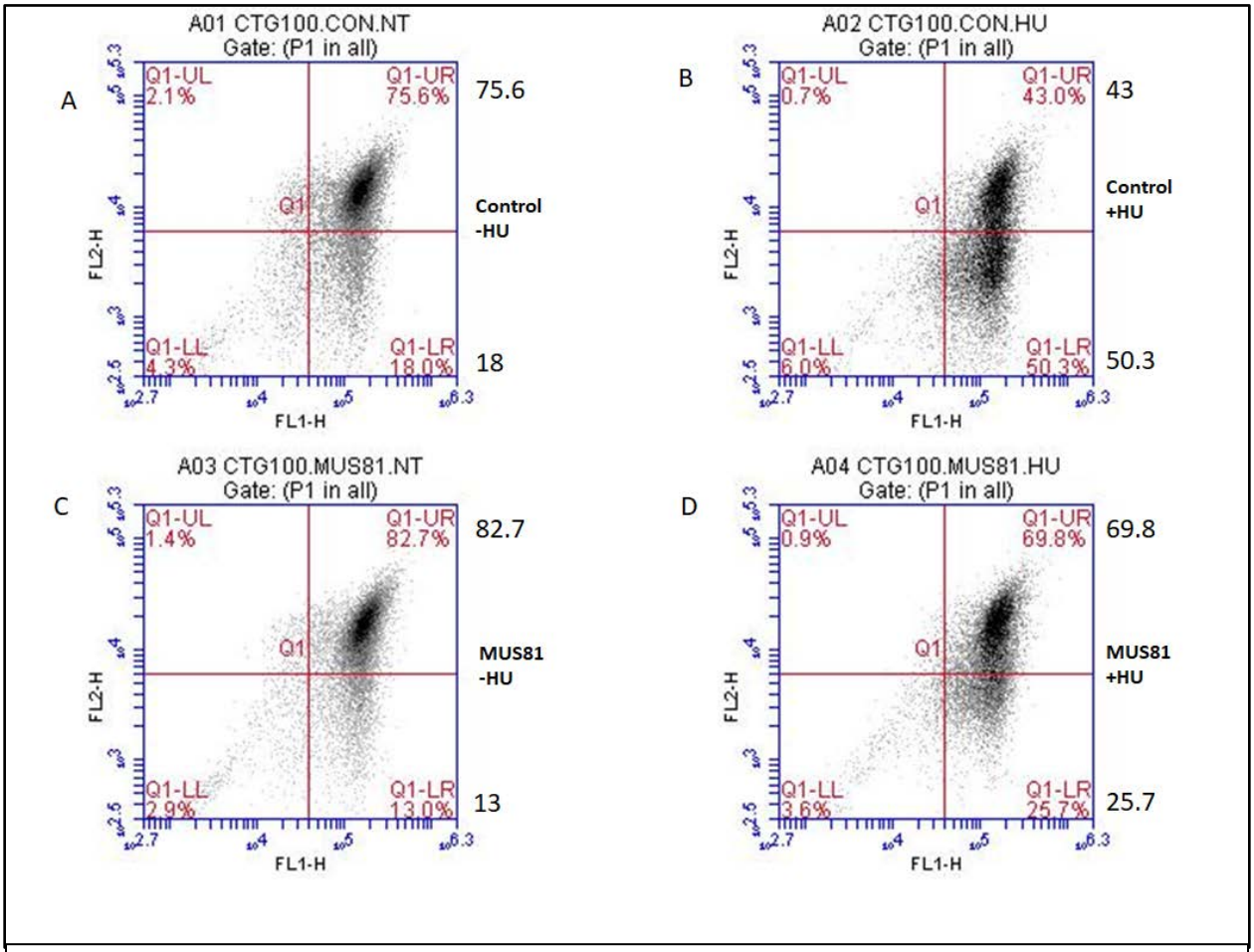



Figure 14. **Loss of MUS81 in CTG100 cells affects DNA stability.** (A) Flow cytometry profile of CTG100 control untreated cells. (B) Flow cytometry profile of CTG100 cells treated with siControl and HU for four days (C) Flow cytometry profile of CTG100 cells treated with siMUS81 (D) Flow cytometry profile of CTG100 cells treated with siMUS81 and HU for four days. siRNA MUS81 knockdown protects against endogenous and HU induced double strand breaks.

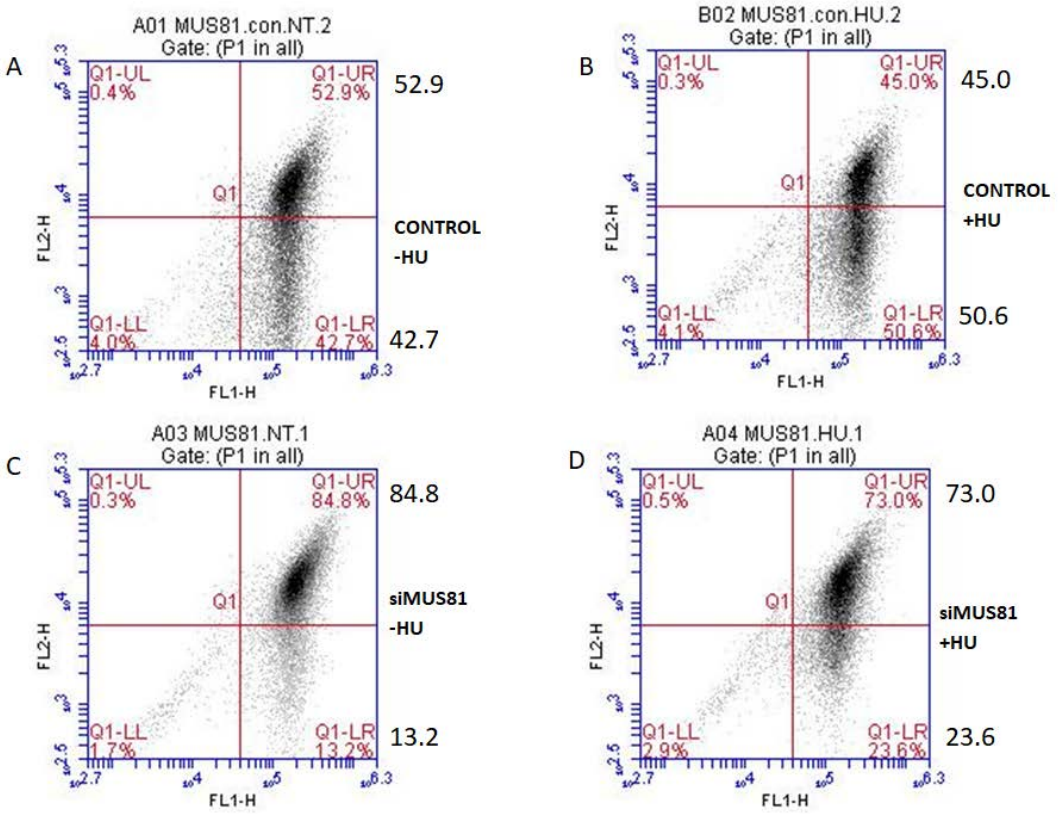



Figure 15. **Loss of MUS81 rescues cells by decreasing events of spontaneous double strand breaks during normal replication.** (A) Flow cytometry profile of CTG100 control untreated cells. (B) Flow cytometry profile of CTG100 cells treated with siControl and HU for four days (C) Flow cytometry profile of CTG100 cells treated with siMUS81 (D) Flow cytometry profile of CTG100 cells treated with siMUS81 and HU for four days. siRNA knockdown of MUS81 protects against endogenous and HU induced double strand breaks.

Figure 16 shows the flow cytometry profile for the $CTG_{100}$ cells treated with siEME1. The untransfected control population in this assay was 75.6% double positive and 18% eGFP. When the control cells were treated with hydroxyurea, the double positive cells decreased to 40.3% and the eGFP cells increased to 50.3%. When treated with siEME1, cells showed a slight increase in double positive cells, increasing to 77.8%, while the eGFP cells decreased slightly to 16.1%. These effects were slightly more pronounced when the assay was performed on cells that were experiencing higher rates of spontaneous double strand breaks, averaging about a 10% increase in double positive cells when treated with siEME1, under normal replication conditions. When treated with siEME1 and hydroxyurea, double positive cells decreased to 47.4% and eGFP cells increased to 45.4%. These results indicate that EME1 is slightly involved in the contributing to microsatellite instability during normal replication process. Knockdown of EME1 does not dramatically protect against double strand breaks.

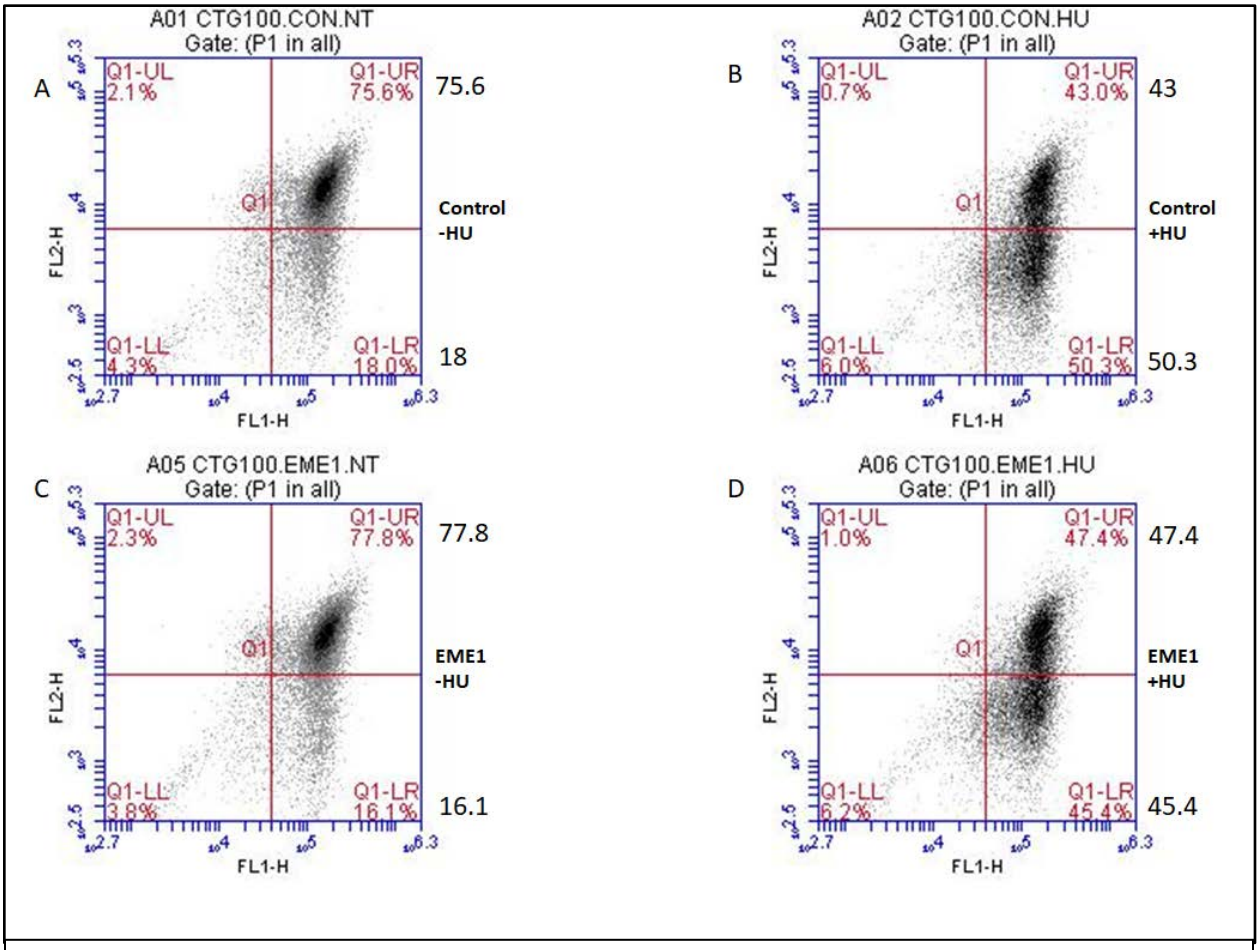



Figure 16. **Loss of EME1 provides minor stability to CTG100 cells.** (A) Flow cytometry profile of CTG100 control untreated cells. (B) Flow cytometry profile of CTG1000 cells treated with siControl and HU for four days. (C) Flow cytometry profile of CTG100 cells treated with siEME1. (D) Flow cytometry profile of CTG100 cells treated with siEME1 and HU for four days. There is slight protection against double strand breaks with EME1 is knocked down.

Figure 17 shows flow cytometry results for siEME2 knockdown assay in the presence or absence of hydroxyurea. The control sample, which showed double positive cells at 75.6%, with eGFP at 18%. When treated with hydroxyurea, double positive decreases to 43.0% and eGFP increases to 50.3%. When treated with siEME2 the double positive cells decrease drastically to 63.6% and eGFP increases to 28.9%. Under hydroxyurea treatment and siEME2, double positive cells decrease further to 38.4% and eGFP increases to 53.5%. Loss of EME2 promotes double strand breaks, possibly by increasing the levels of MUS81-EME1 dimers. In contrast, knockdown of EME2 neutralizes the S-phase effect of hydroxyurea on double strand breaks. This assay shows that knocking down EME2 causes the DF2 Myc CTG100 cells to alter their ability to tolerate non-B DNA structures formed from the CTG/CAG repeat. EME2 is an important protein to the cell during DNA replication under normal conditions, and important in the crossover junction endonuclease function of the cell during replication stress.

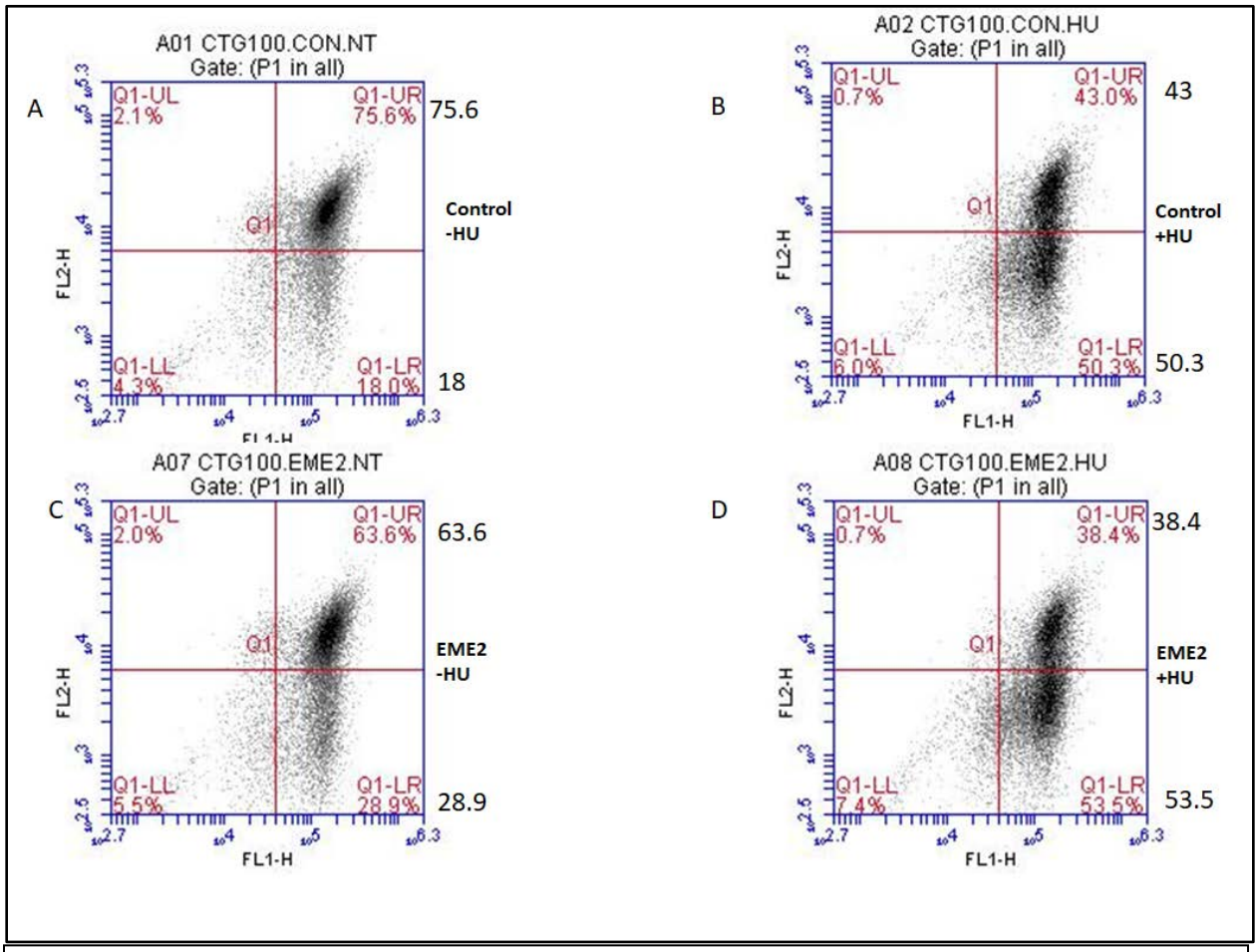



Figure 17. **Loss of EME2 in CTG100 cells affects DNA damage tolerance.** (A) Flow cytometry profile of CTG100 control untreated cells. (B) Flow cytometry profile of CTG100 cells treated with siControl and HU for four days (C) Flow cytometry profile of CTG100 cells treated with siEME2 (D) Flow cytometry profile of CTG100 cells treated with siEME2 and HU for four days. siEME2 enhances endogenous breaks. The knockdown of EME2 does not dramatically reduce HU effect.

Figure 18 shows a summary of the statistical analysis acquired from the siRNA treatment of MUS81, EME1, and EME2 in the CTG100 cells. Using standard normal value equation $z = \frac{x - \mu}{\sigma}$ where x is the single sample value, μ is the mean, and σ is the standard deviation. P value was calculated from the z score using one-tailed hypothesis, and a significance level of 0.05. The bar graph shows the mean for the particular data, including all the control values for that particular category plus the single sample value. The standard deviation was calculated with the same data that was included in the bar graph. A p value $< 0.05$ is considered significant. The pattern observed with the CTG100 cells treated with siMUS81 showed statistical significance when compared with the control both in the no treatment (NT) and HU samples, indicating a consistent observed increase in amount of yellow cells detected by flow cytometry. Also of significance was the consistent decrease in amount of yellow cells when CTG100 cells were treated with siEME2 during normal replication.

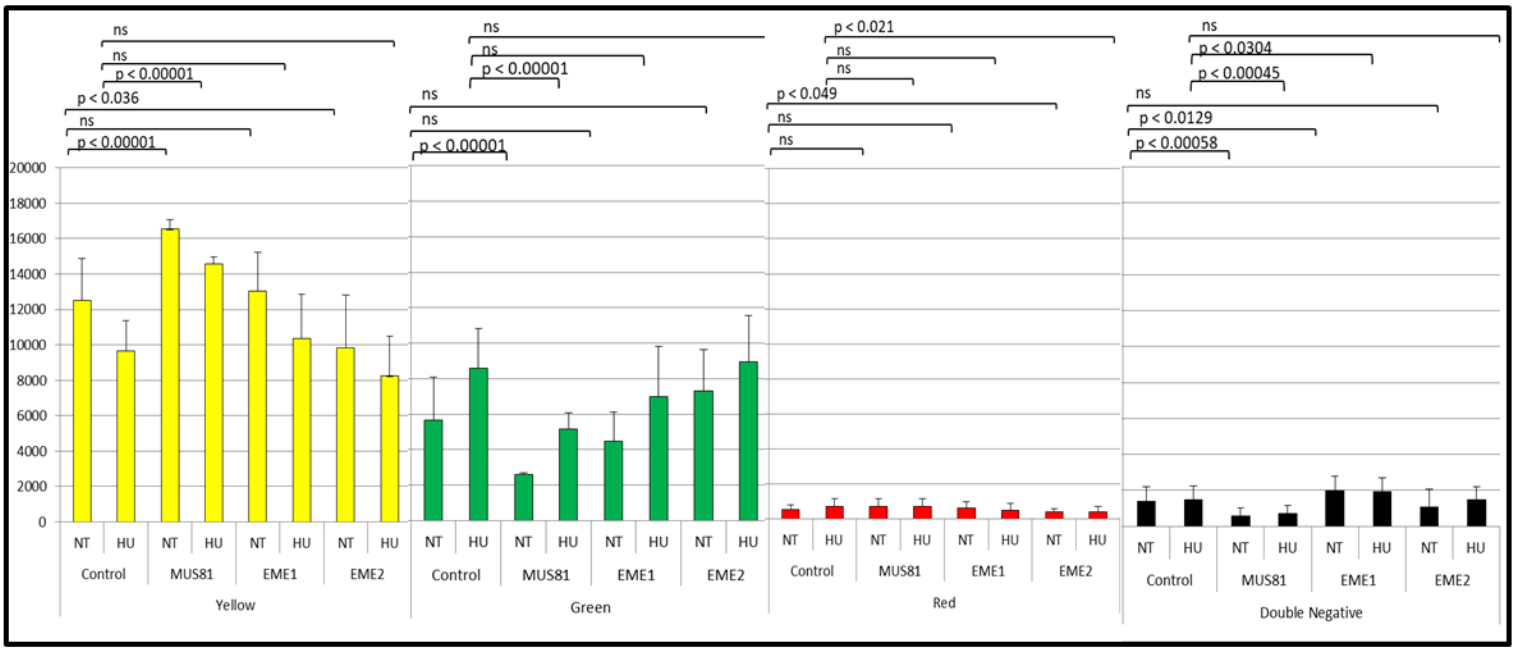



Figure 18. **Statistical analysis of the flow cytometry results from siMUS81, siEME1, and siEME2.** Analysis was derived from cellular populations of each of the repeated assays with the various siRNA treatments. The p-value shows significance of the compared numbers. The standard deviation bar shows the average range of the compared populations. Discrepancy between the statistics shown in the graph compared to flow cytometry data is caused by the variability between repeats of the experiments.

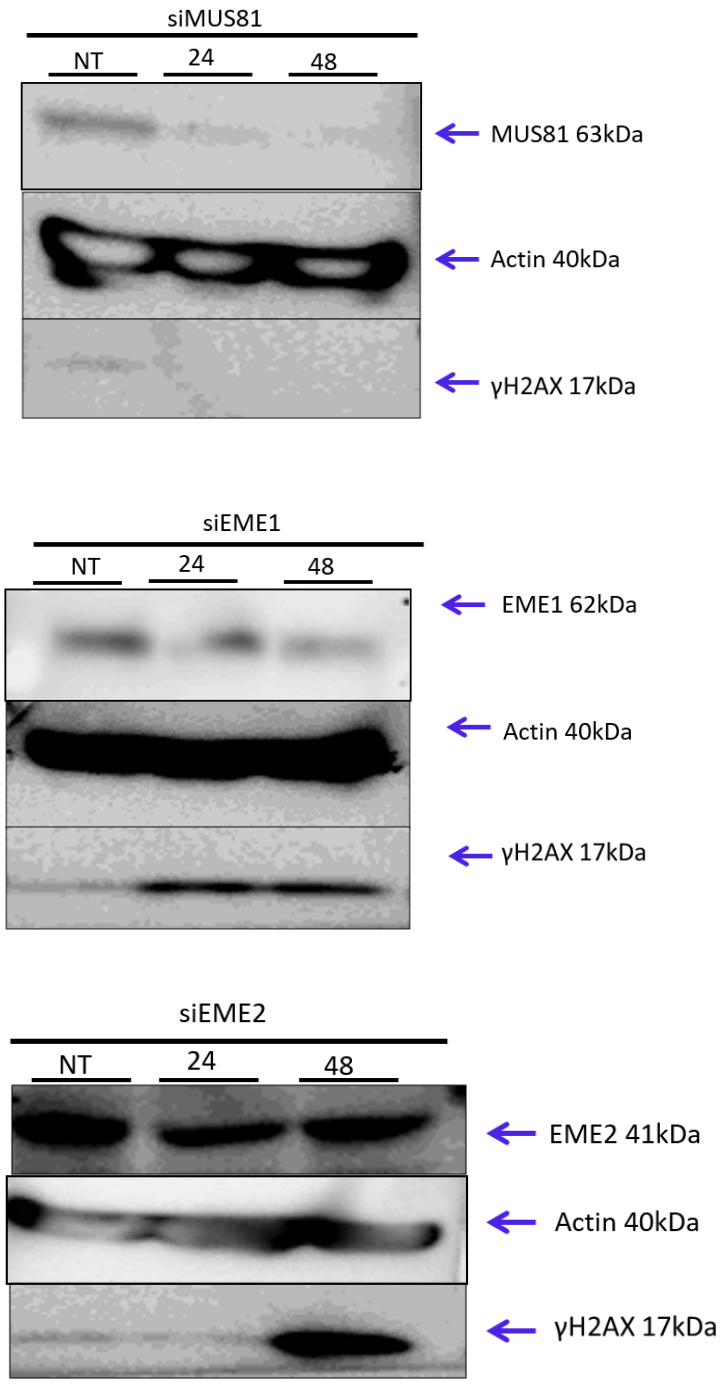



Figure 19. **Western blot proof of concept siRNAs.** Western blot showing the effect of siRNA treatment on CTG23 cells for siMUS81 (A), siEME1 (B), and siEME2 (C), cells were collected at time points 24 and 48 hours after treatment as well as no treatment (NT) cells for control. γH2AX levels were also measured to see if knocking down of the proteins mentioned previously caused replication stress and DSB. MUS81 knockdown caused a decrease in basal γH2AX signal. EME1 knockdown caused a slight increase in γH2AX signal at 24 and 48 hours, and knockdown of EME2 caused a large increase in γH2AX signal at 48 hours.

## V. DF2 Myc $CTG_{100}$ cells spPCR of ectopic and endogenous microsatellite

Since the flow cytometry showed a noticeable change in the CTG100 cells caused by the knockdown of MUS81, EME1 and EME2, we wanted to see if there were any noticeable breaks or changes within the DNA, namely with PCR. First we checked that the siRNA successfully knocked down its target protein with Western blot as seen in Figure 19.

We wanted to check the expanded microsatellite, as well as an endogenous microsatellite, in this case the DMPK locus. DMPK is also a CTG microsatellite, so we thought it would be a good comparison for the expanded ectopic $CTG_{100}$ microsatellite. The PCRs shown here are of the population of DF2 Myc $CTG_{100}$ cells after treatment with siRNA with or without hydroxyurea. The PCR primers work by having primers attach on either side of the ectopic site, or either side of the endogenous DMPK locus. Also included in each PCR were primers that attach elsewhere in the genome, not associated with any microsatellite, to act as a loading control. The ectopic microsatellite samples, as well as the endogenous DMPK samples contained 50 pg of genomic DNA (small pool PCR).

Despite the changes seen on the flow profiles after purifying the genomic DNA, there are few if any noticeable differences on the polyacrylamide gels. This is apparent in both the gels that show the ectopic site as well as the gels that show the DMPK site. This is probably because even though the flow profiles show change, the double positive cells still number about half of the total cells even after siRNA treatment, or treatment with

hydroxyurea. This does not mean that there is no genomic change, it just means that what changes have occurred are few and cannot be seen with this method (Figure 20, and 21).

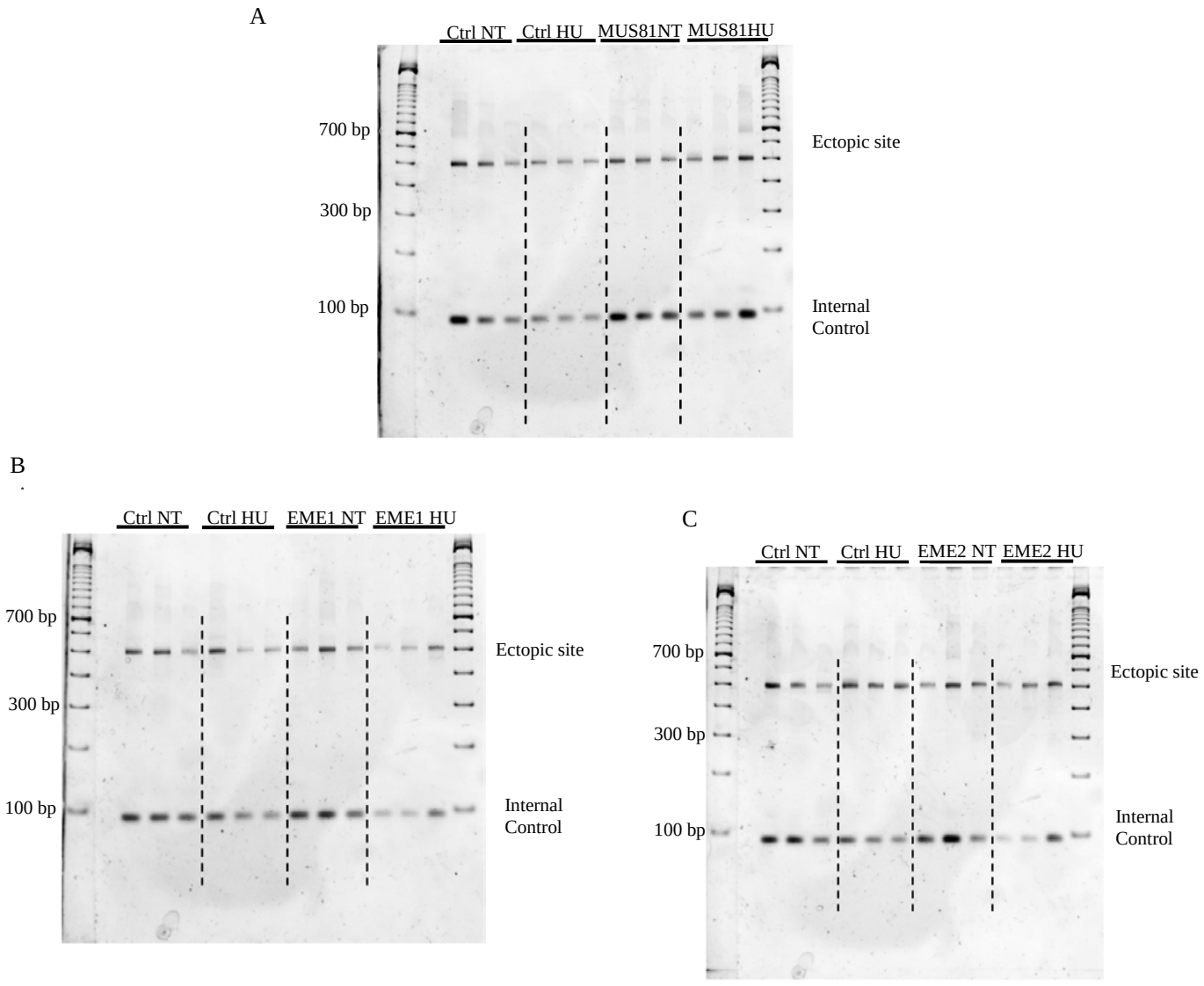



Figure 20. **Small pool PCR of CTG100 cells treated with various siRNAs and HU show slight loss of microsatellite signal.** (A) spPCR of siMUS81 treated CTG100 cells. (B) spPCR of siEME1 treated CTG100 cells. (C) spPCR of siEME2 treated CTG100 cells.

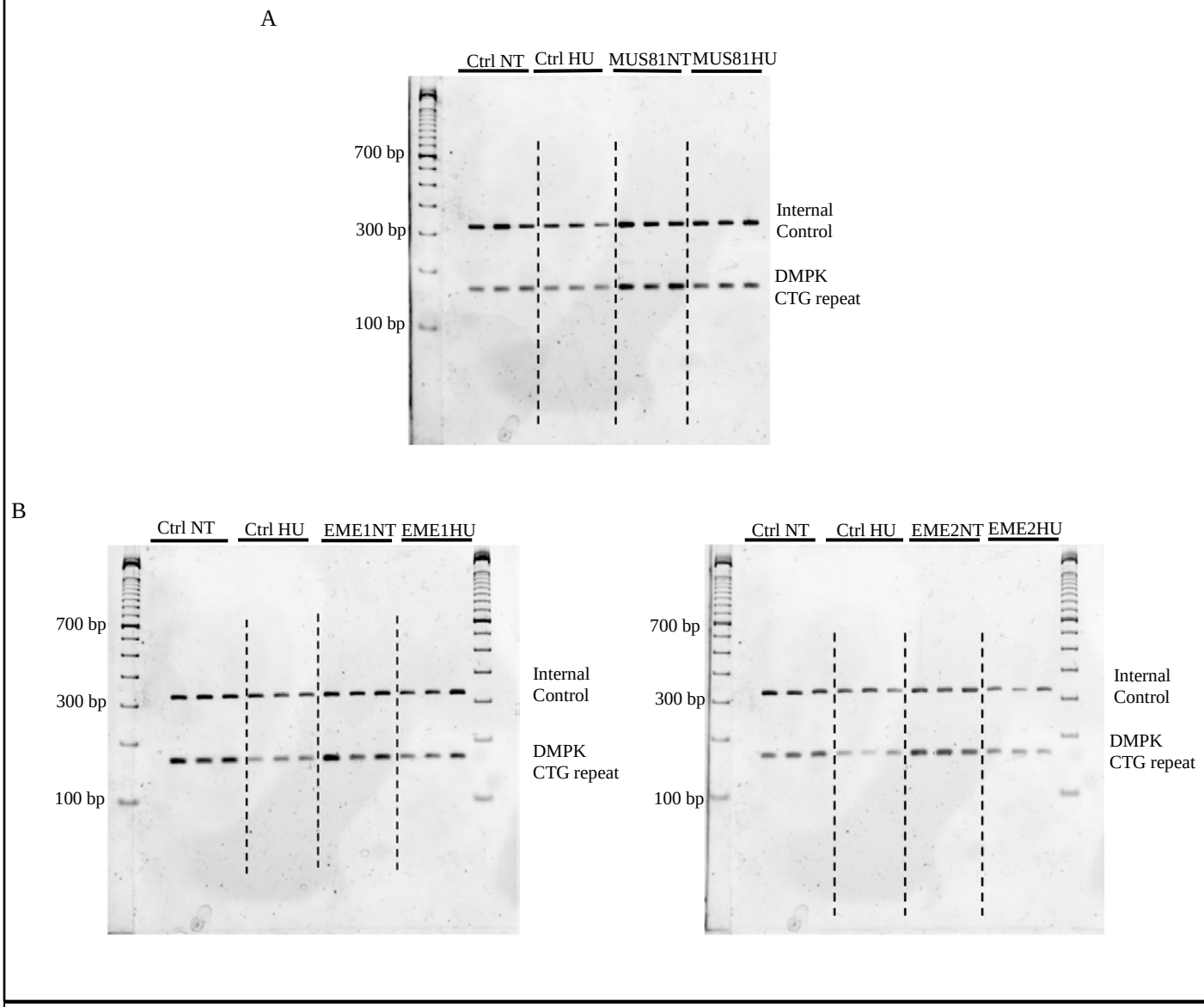



Figure 21. **CTG100 cells treated with siRNA and HU show no loss of signal for endogenous microsatellite DMPK.** (A) siMUS81 spPCR of endogenous DMPK. (B) siEME1 spPCR of endogenous DMPK microsatellite. (C) siEME2 spPCR of endogenous DMPK microsatellite.

## VI. DF2 Myc $CTG_{100}$ cells sorted by color marker

In order to perform more sensitive testing for mutations caused by siRNA and hydroxyurea treatments, DF2 Myc $CTG_{100}$ cells were taken to Cincinnati Children's Hospital and sorted according to color marker using their five laser Arni flow cytometer. The $CTG_{100}$ cells were sorted according to color, either double positive (yellow, red+ green), eGFP (green), or double negative (no color markers). The samples sorted were cells pre-treated with siMUS81, siEME1, or siEME2, all of which had been treated with hydroxyurea for four days and given four days to recover. The four days of recovery is important because it allows for degradation of the color marker proteins from the start of the assay. This will show only the newly synthesized color marker proteins, which could change depending on what effect the assay had on the template. This recovery time allows for the separation of the color markers. Around 20,000 cells from each sample were sorted and used for spPCR.

PCR and spPCR are different in terms of detecting changes in a very small population of cells. PCR that was performed on cells sorted compared to the unsorted cells gave different results. This is probably because even though spPCR uses a miniscule amount of template (50 pg), usually 50% of the cells were still double positive with double negative cells usually being about 10% of the total sample and GFP cells around 30-40% of the total population. Even with using 50 pg of sample the likelihood of selecting template that did not have double strand breaks was still more likely. Cell sorting eliminated this probability and allowed the changes that occurred because of the

various protein knockdowns to be clearly seen. The results that show the outcome from the cell sorting are Figures 22-24.

As in the previous experiment, we wanted to look for any expansions, contractions, or breaks, within the expanded microsatellite of the ectopic site as well as at the endogenous DMPK locus. With so little sample to work with we combined the primers for the ectopic site and the DMPK locus in each spPCR. In Figure 22, the double negative samples for both MUS81 and EME1 gave very contracted bands (*) arising from slipped strand reannealing during PCR. This shows the loss of ectopic signal relative to DMPK in several lanes, as well as the very short PCR products in 22A and 22B. The short products seen in 22A and 22B are likely PCR products of loss of the CTG repeat, as well as loss of flanking DNA.

Figure 22C however, shows the most interesting and unexpected results. The double negative samples show loss of the ectopic microsatellite signals, as well as contraction of the DMPK locus in a fraction of the cells (also seen in Figure 23B). Knockdown of EME1, MUS81 and EME2 all gave changes at the ectopic site and at the endogenous DMPK in a subpopulation of cells (double negative). These data suggest that MUS81, EME1 and EME2 have prominent roles in maintaining stability in both larger and smaller microsatellites; as shown by the effect these knockdowns had in both the exogenous and endogenous microsatellite repeat.

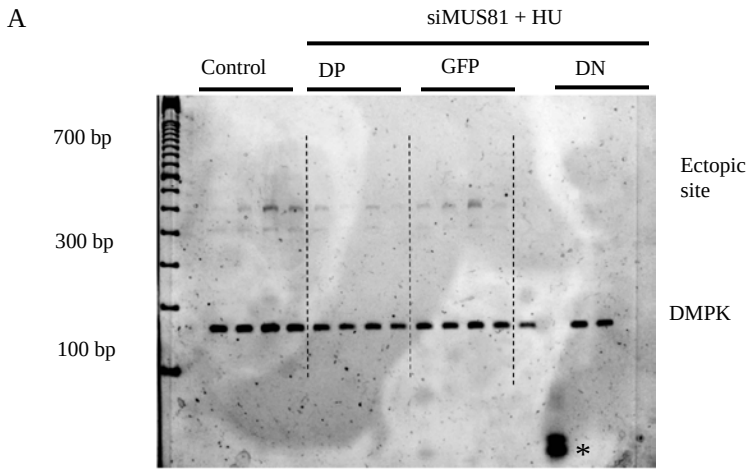
We wanted to explore the possible mutation within the endogenous DMPK locus, so the spPCR was repeated with the DMPK primers as well as traditional Internal Control

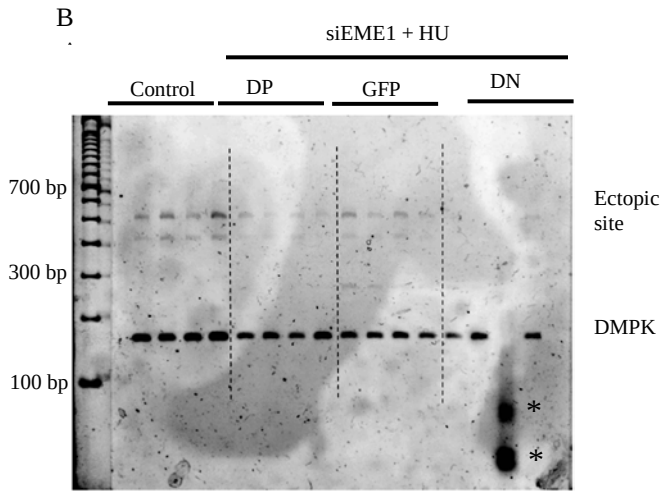
primers, which bind elsewhere in the genome and are not associated with any microsatellite. This is seen in Figures 23 and 24. Figure 23 A and B show DMPK locus of siMUS81 and siEME1 treated cells respectively. The MUS81 treated cells showed no change within the DMPK locus. Figure 23 B siEME1 samples show a new band in double positive cells treated with hydroxyurea. This suggests that a small population of cells had a small microsatellite contraction occur without loss of any color marker. Figure 24 A shows the polyacrylamide gel of the double positive sample and the eGFP sample of the DMPK microsatellite. The faint bands above the dark DMPK band indicate some expansion occurring within the endogenous microsatellite. Figure 24 B shows the double negative sample of the siEME2 treated cells. The DMPK shows primarily contraction of the DMPK locus, but also several expansion events with a band between the primary DMPK band and the Internal Control band that does not appear on the Control portion of the gel. The new bands that appear in Figure 24 B lanes 5-8 are consistent with spontaneous breaks seen by siEME2 treated cells in Figure 17. This is consistent with MUS81-EME2 activity during S-phase of the cell cycle.

Cells were sorted according to change in color marker at the expanded microsatellite within the ectopic site (double positive, green, and double negative). However the double negative cells show deletion and instability at an unrelated DMPK site. In cells that are affected by the siRNA both ectopic and endogenous repeats are affected. This indicates that the knockdown of EME2, in combination with replication

stress, leads to microsatellite expansion and contraction not only of expanded microsatellites but also of endogenous microsatellites within the cell.

These results suggest that a subpopulation of cells is hypersensitive to DSBs following EME2 knockdown plus HU; this hypersensitivity is reflected at the ectopic and endogenous CTG microsatellites.



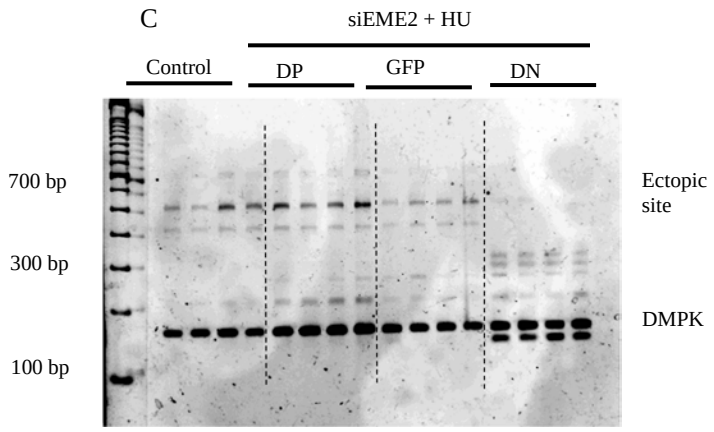




Figure 22. **Sorting cells by color shows mutations occurring because of siRNA knockdown.** (A) siMUS81 spPCR showing ectopic site and endogenous DMPK. (B) siEME1 spPCR showing ectopic site and endogenous DMPK. (C) siEMEM2 showing ectopic site and endogenous DMPK. All knockdowns show effects at DMPK and the expanded ectopic microsatellite region in a fraction of double negative cells.

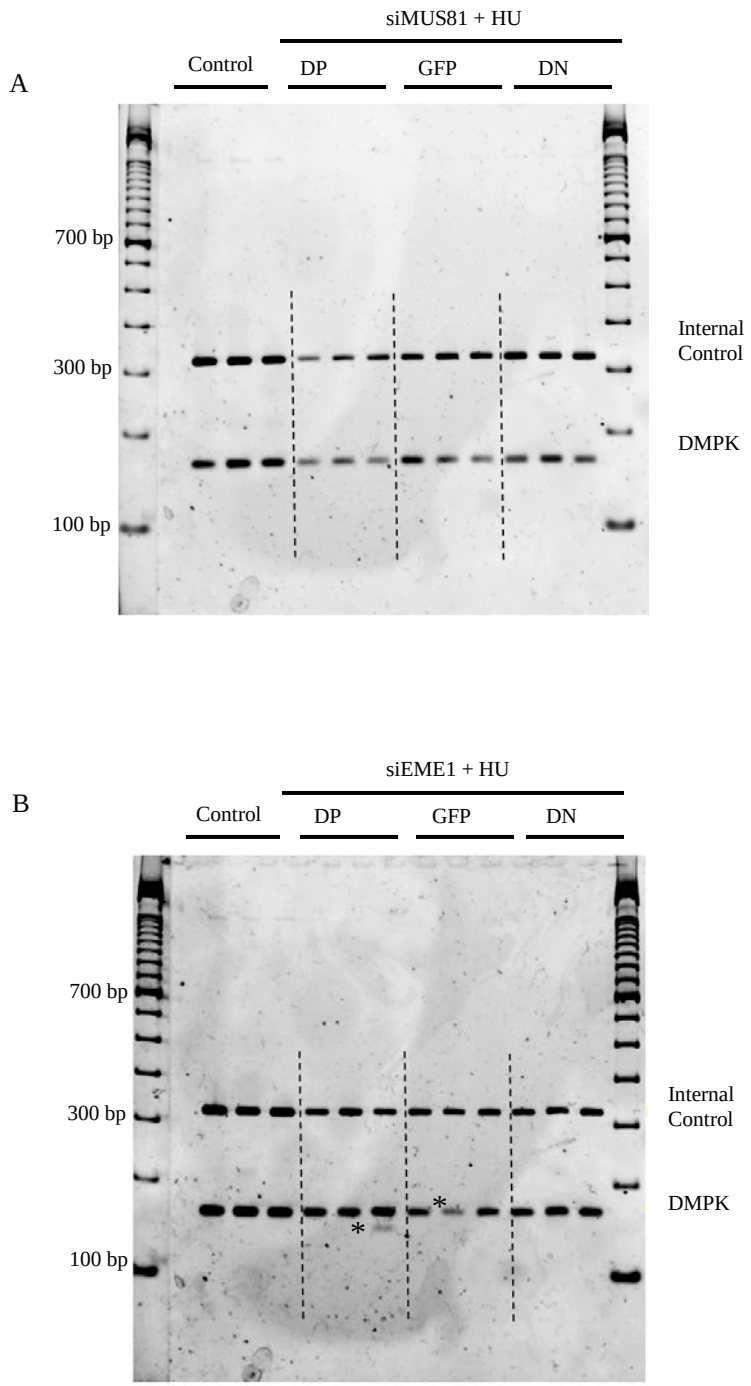



Figure 23. **siMUS81 and siEME1 have slight effect on endogenous microsatellite DMPK.** (A) CTG100 cells treated with siMUS81 separated by color markers spPCR of DMPK. (B) CTG100 cells treated with siEME1 separated by color markers spPCR of DMPK. Samples contain 50 pg of template each. This conclusion is different to the conclusion form fig 25 A and B. This could be because of the diversity of spPCR sampling, or due to knockdown efficiency. siEME1 samples showed slight contraction within the endogenous microsatellite in the double positive sample and slight loss of signal in the GFP sample, indicating presence of DSBs.

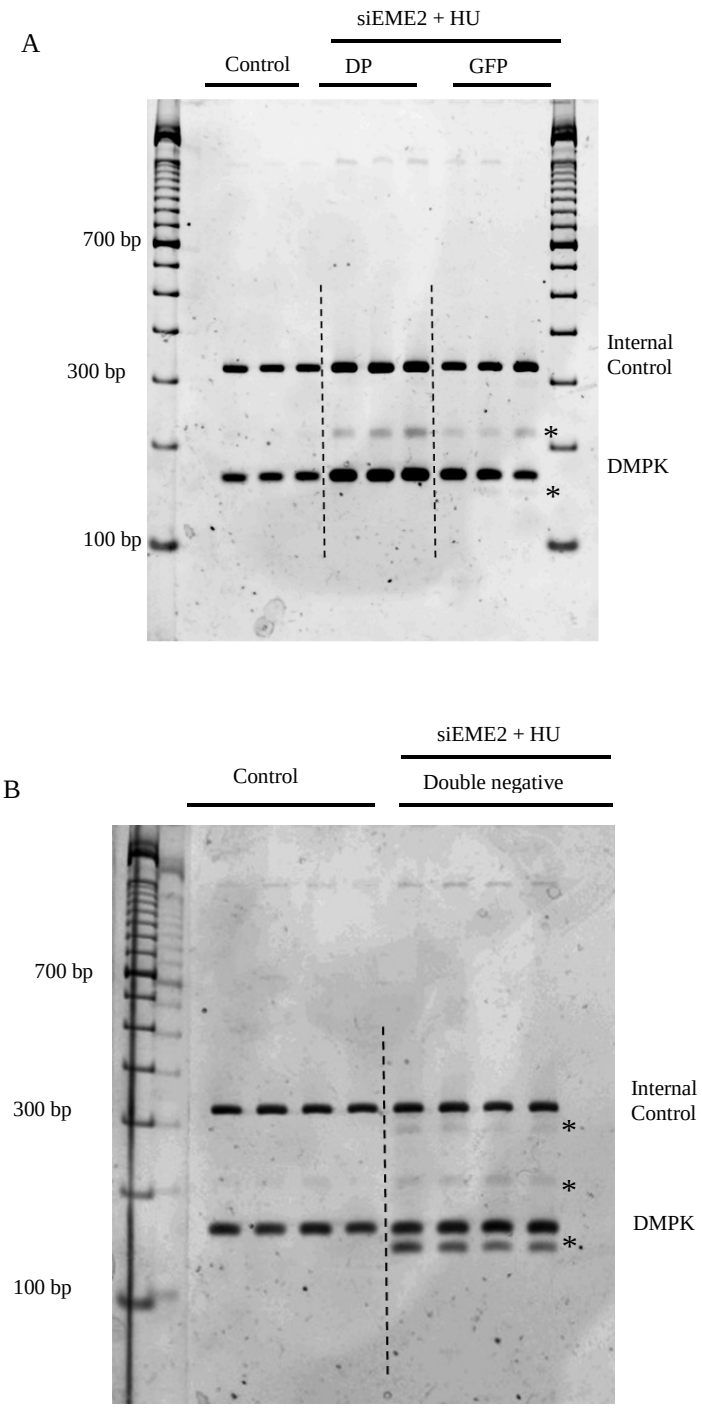



Figure 24. **siEME2 treatment leads to contraction of endogenous microsatellite DMPK.** (A) CTG100 cells treated with siEME2 sorted by color marker double positive (yellow) and GFP (green) spPCR for DMPK. (B) CTG100 cells treated with siEME2 sorted by color marker double negative (loss of both dTomato and eGFP) spPCR for endogenous DMPK. While knockdown of EME1 seems to cause some mutations, it does not have a higher mutation frequency to siEME2.

## VII. **DF2 Myc $CTG_{100}$ sorted red cells after treatment with shRNA and siMUS81**

A striking consistent observation was that a significant percentage of cells DF2 Myc $(CTG)_{100}$ cells treated with I-Sce1 turned red (Figure 25 E-H), whereas the same cells spontaneously turned green, which was exacerbated by treatment with HU (Figure 25 A-D). These results suggest that a "clean" DSB generated by I-Sce1 restriction enzyme digestion could be processed by cells to allow homologous recombination between the Alu sites flanking the eGFP gene, but that a replication-dependent double strand break could not be processed this way.

To begin to identify which proteins could be involved in the difference of color markers seen from cells treated with I-Sce1, and cells that normally turn green. Samples were taken to Cincinnati Children's Hospital to use the flow cytometer, and sorted by dTomato (red) color markers. These samples were the DF2 Myc $CTG_{100}$ cells transfected with a mixture of about fifty shRNA plasmids, and treated with hydroxyurea for four days, and allowed to recover for four days before being sorted. Also sorted for red cells were siMUS81 treated cells without hydroxyurea treatment. The results of this treatment and sorting are seen in Figure 26.

As shown in previous work in the thesis by Rujuta Gadgil (2016), eGFP positive cells are unstable and die within about a week of shifting to green. This is because of the position of the trinucleotide repeat. The centromere is to the right (Figure 6) of the CTG trinucleotide repeat. If DNA undergoes a double strand break that cannot be repaired, it

leads to cellular death. This break is catastrophic to the cell because it is a significant loss of the dTomato region and beyond is a significant portion of the arm of the chromosome. Because this loss is so extensive it may cause the positive GFP cells to undergo apoptosis causing the GFP cells to gradually die off and the flow profile to shift back from green to yellow (Gadgil 2016).

The chromosome upstream of the eGFP site has a catastrophic double strand break occur due to replication stress. The loss of the acentric chromosome fragment that transpires is not sustainable and the cells that this happens to in die within days of it occurring. Theoretically dTomato positive cells should have a stable recombination, and therefore, were allowed to set down on a 10 cm plate and grow to confluency. After reaching confluency on the plate, cells were collected and flow cytometry profiles were taken for the shRNA mixture of cells and the siMUS81 red cells seen in Figure 26 A, and B respectively. The shRNA treated cells show 35.8% dTomato positive and 57.7% double negative. And the siMUS81 treated cells show 55.1% dTomato positive and 33.5% double negative.

After completion of flow cytometry, the sample templates were purified for PCR and spPCR. Figure 27 A shows the spPCR results of the ectopic site of the isolated red cells. The spPCR showed new bands above ectopic band in siMUS81 red cells in the final lane of the siMUS81 section, as well as the middle and last lane of the shRNA treated mixture red cells.

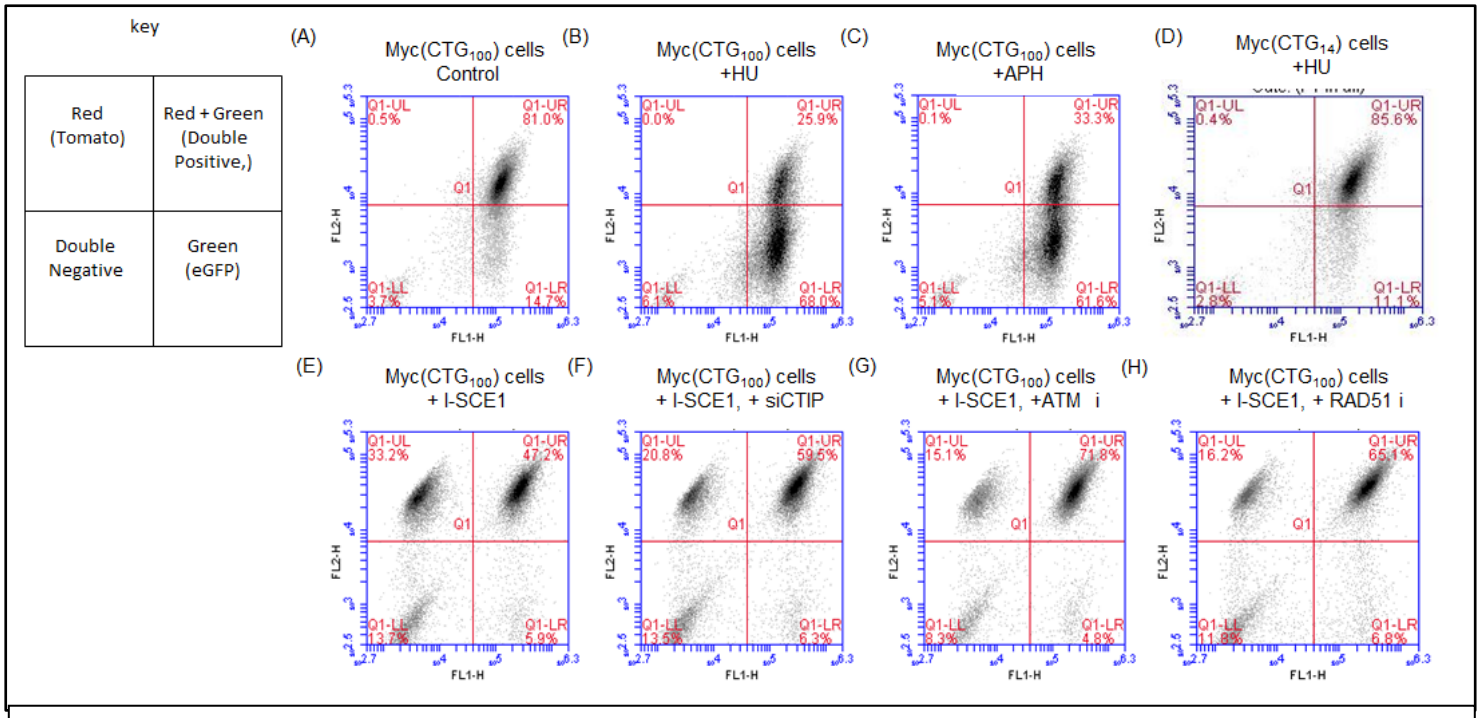



Figure 25. **Differential Processing of DNA DSBs in CTG100 cells.** Depending on the type of break that occurs determines what color the cells shift to. (E-H) When cut with the enzyme I-SCE1 a clean double strand break occurs and cells shift to red. (A-D) When a double strand break occurs due to replication stress brought on by various drug treatments (HU, or APH) cells shift to green. This does not occur in CTG14 cells because it does not have an expanded microsatellite. Figure provided by Dr. Todd Lewis.

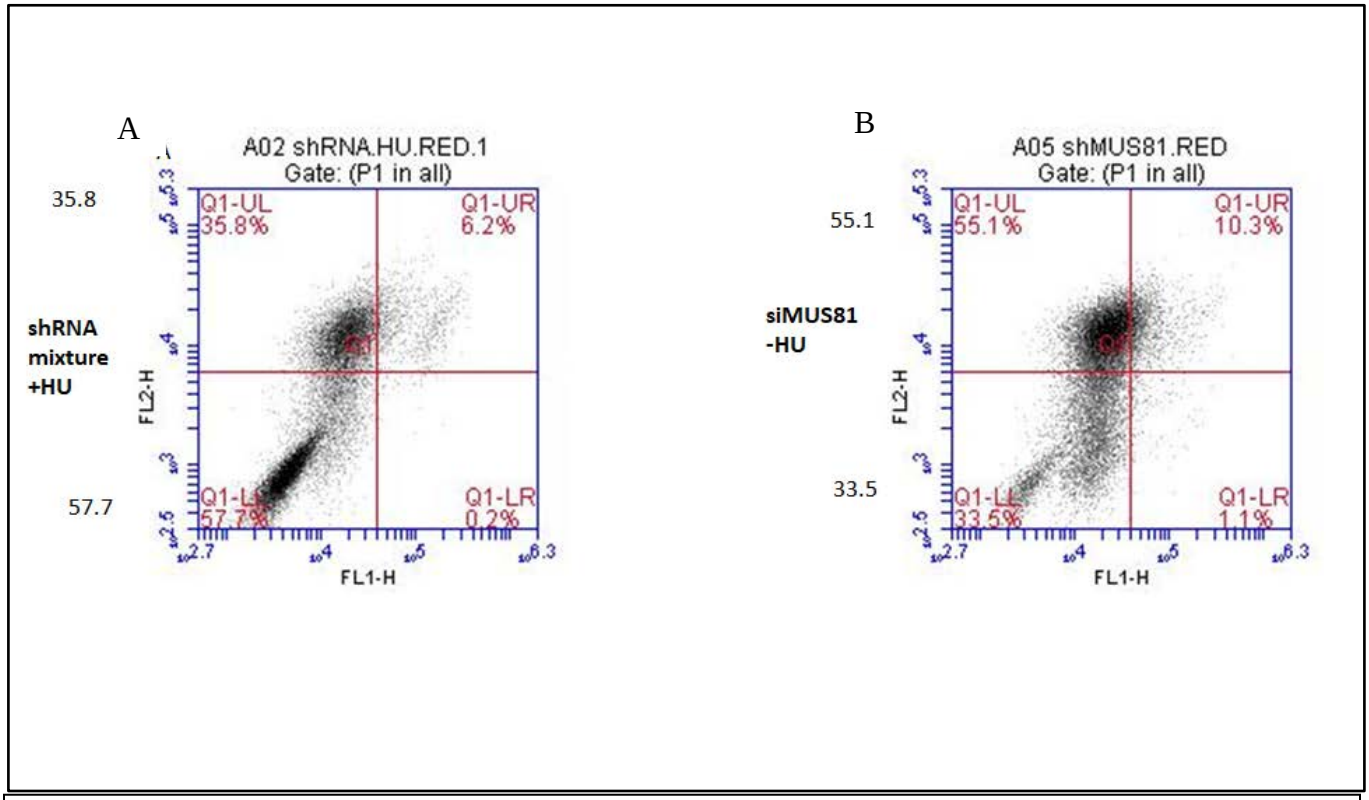



Figure 26. **Cell sorted shRNA treated dTomato cells flow cytometry.** Flow cytometry results for (A) shRNA mixture treated CTG100 cells treated with HU for four days and sorted for red cells by flow cytometry. (B) CTG100 cells treated with siMUS81 without HU and sorted for red cells by flow cytometry. Flow cytometry profile indicates that red cells continue to undergo loss of dTomato.

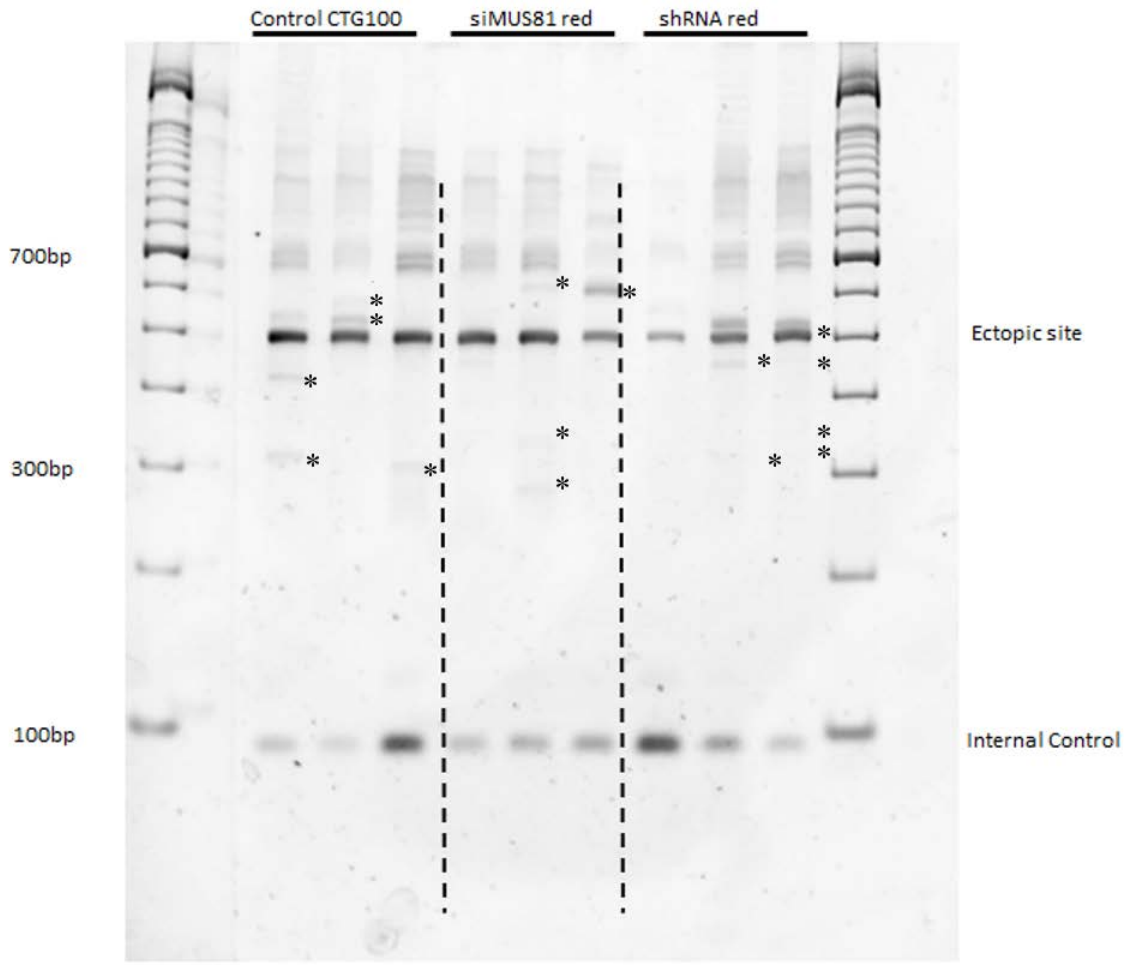



Figure 27. **Cell sorted shRNA treated dTomato cells PCR.** spPCR of ectopic site from sorted red siMUS81 and shRNA mixture cells. Samples contained 50 pg template The (*) seen in this figure of the dTomato positive cells is the various expansion and contraction events that have occurred in the isolated red cells. The fading seen in the first lane of the shRNA red sample is indicative of loss of ectopic site signal, or double strand breaks. The majority of the red positive cells seem to be point mutations within the eGFP site causing loss of the green protein making the cells appear red, as indicated by the ectopic site band appearing primarily at the same length as the CTG100 control cells.

## VIII. **DF2 Myc $CTG_{23}$ Cells treated with siRNA showed change is cell cycle**

MUS81 forms complexes with either EME1 or EME2, depending on the phase of the cell cycle. In order to see if the siRNA treatment had an effect on the cell cycle DF2 Myc $CTG_{23}$, cells were treated with siMUS81, siEME1 or siEME2 and collected at 24 or 48 hours after treatment, along with a control treated with siControl collected 48 hours after treatment. Samples were treated with DRAQ7 dye (Cell Signaling) and analyzed by flow cytometry. The results of this assay are seen in Figure 28.

Figure 28 A shows siControl treated cells, which exhibits the usual G1-S-G2-M phases. The siMUS81 treated cells collected at 24 hours (Figure 28 B) show an increase in cells in S phase. At 48 hours (Figure 28 C) the first peak, indicating G1 phase, has decreased which could indicate a slowing of cell cycle progression as not as many cells are cycling back to the G1 phase. The large sub G1 phase could indicate a population of cells undergoing apoptosis as a result of the siRNA treatment.

The cells treated with either siEME1, or siEME2 have a dramatic shift in the DRAQ7 charts, when collected at 24 hours (Figure 28 D, and F). The siRNA treatment seems to have caused cell cycle arrest in S phase. After 48 hours EME1 treated cell profile looks like the control and has returned to normal (Figure 28 E). EME2 cells after 48 hours are returning to normal with the reappearance of the G1 peak but the S-G2 peaks have not recovered completely. Knocking down of MUS81, EME1, and EME2 has effects on how cell cycle is able to progress; slowing progression with depletion of MUS81 and S-phase arrest with depletion of EME1, and EME2.

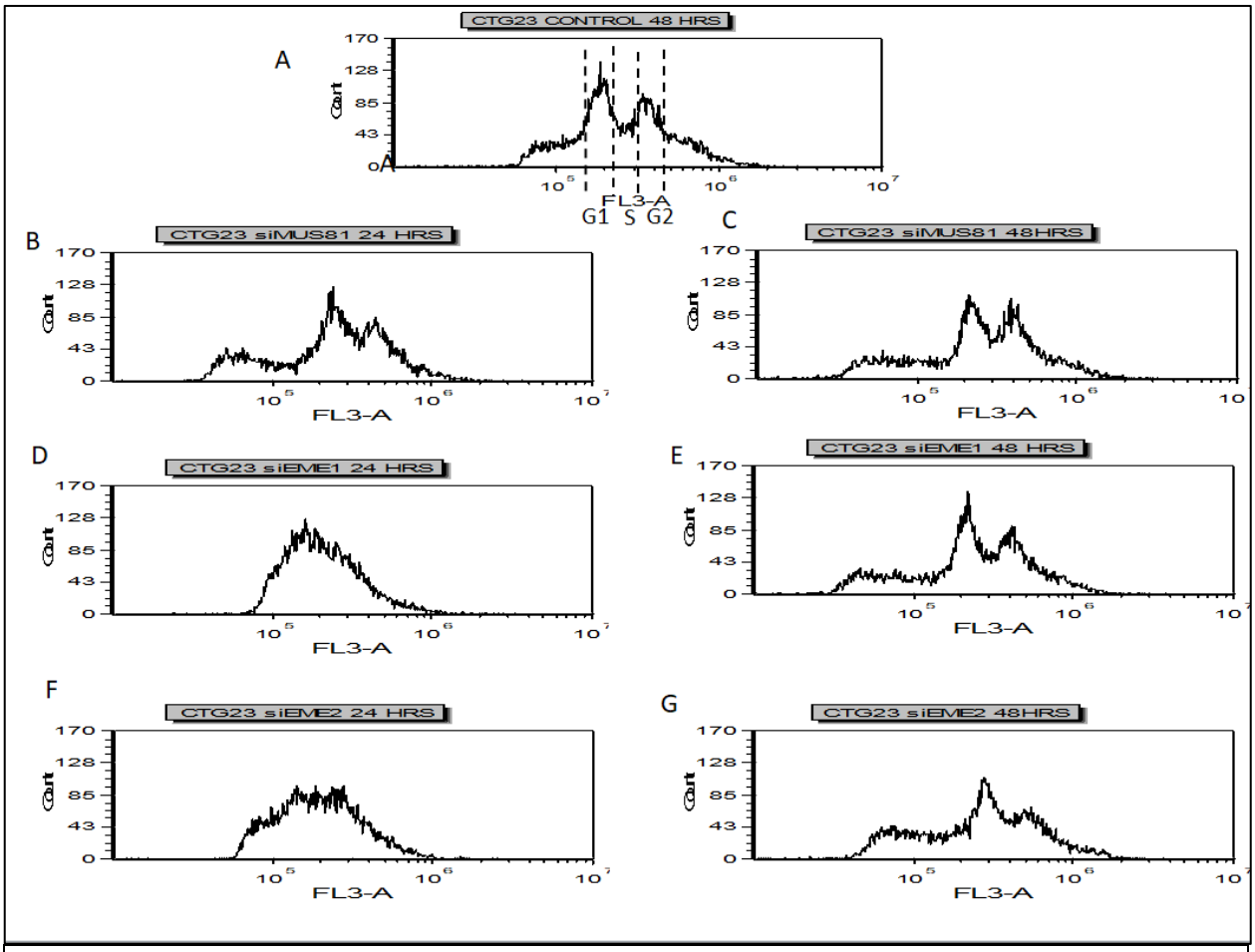



Figure 28. **Treatment with siMUS81, siEME1, and siEME2 cause change in cell cycle.** (A) Control CTG23 cells treated with siControl. (B) CTG23 cells treated with siMUS81 collected 24 hours after transfection. (C) CTG23 cells treated with siMUS81 48 hours after transfection. (D) CTG23 cells treated withsiEME1 collected 24 hours after transfection. (E) CTG23 cells treated with siEME1 48 hours after transfection. (F) CTG23 cells treated with siEME2 collected 24 hours after transfection. (G) CTG23 cells treated with siEME2 collected 48 hours after transfection. Treatment with siMUS81 at 24 hours caused a shift in the x-axis. Treatment with siEME1 and siEME2 caused an S phase arrest of the cells.

## IX. **DF2 Myc $CTG_{23}$ cells show no change in flow cytometry profile in response to siRNA treatment.**

As additional controls for the siRNA treatments and results seen within the CTG100 cells, namely the expansion and contraction of the endogenous microsatellite DMPK, we thought it important to see if knocking down of MUS81, EME1, and EME2 had any effect on the CTG23 cell lines, which contain a wild type length CTG microsatellite at the ectopic site. The CTG23 cells were treated with the aforementioned siRNAs with or without hydroxyurea for four days, and allowed four days to recover from treatment. These results can be seen in Figure 29-31.

The siControl flow profile showed 94.5% double positive cells and 0.7% eGFP positive cells. After treatment with hydroxyurea, the numbers remain unchanged. This has been demonstrated in previous studies and this cell line is more stable and resistant to replication stress. After treatment with the respective siRNAs with or without hydroxyurea, the flow cytometry charts remain mostly unchanged. The CTG23 cell line is more stable and more resistant to double strand breaks than the CTG100 cell line. This is due to fewer repeats in the CTG23 cells compared to the CTG100 cells. This offers the cells more stability, despite the changes in cell cycle at 24 and 48 hours in the DRAQ7 assay, compared to the flow results of hydroxyurea vs. recovery.

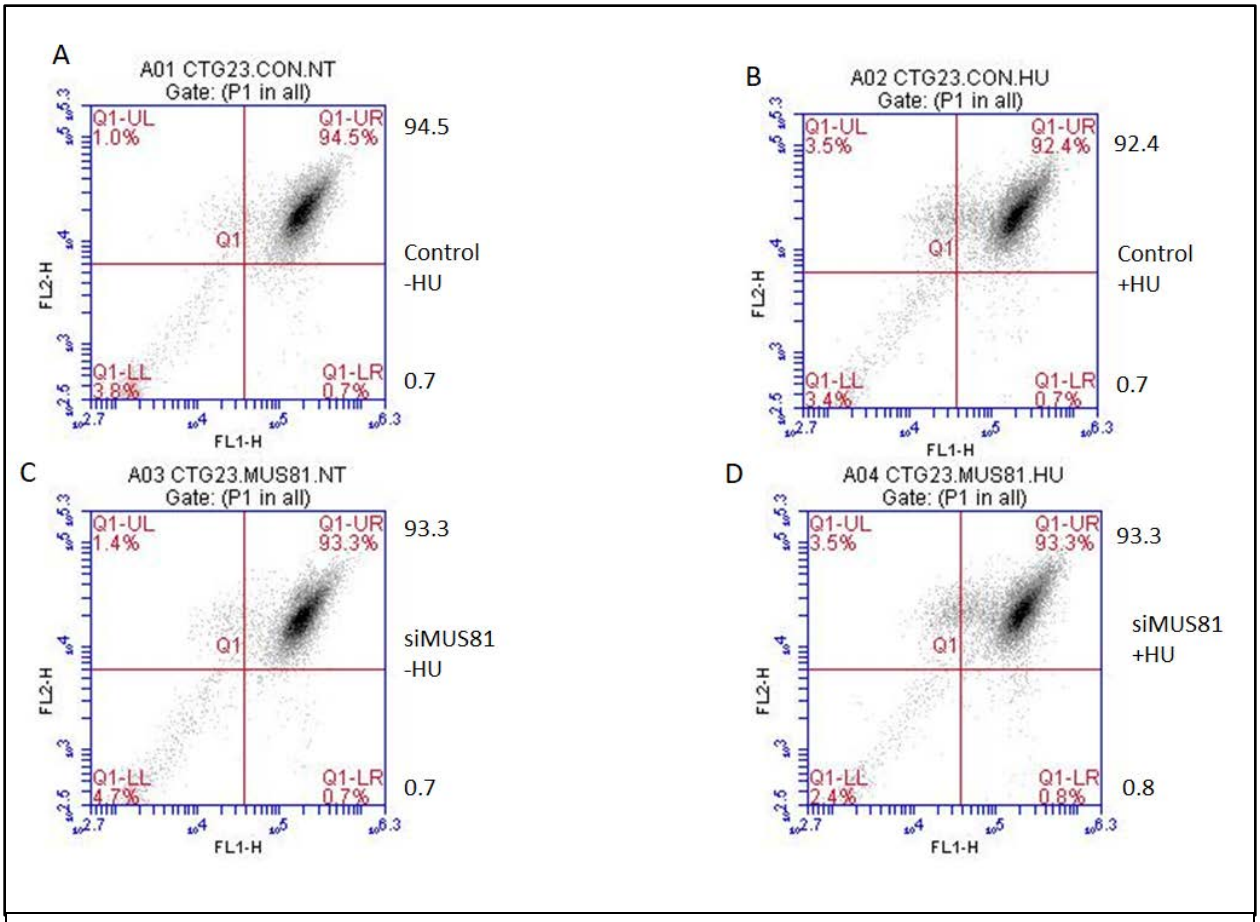



Figure 29. **siMUS81 treatment has no effect on DSBs in CTG23 cells.** (A) CTG23 cells control with no treatment of HU (B) CTG23 control cells with HU treatment. (C) CTG23 cells treated with siRNAMUS81 without HU. (D) CTG23 cells treated with siMUS81 and HU.

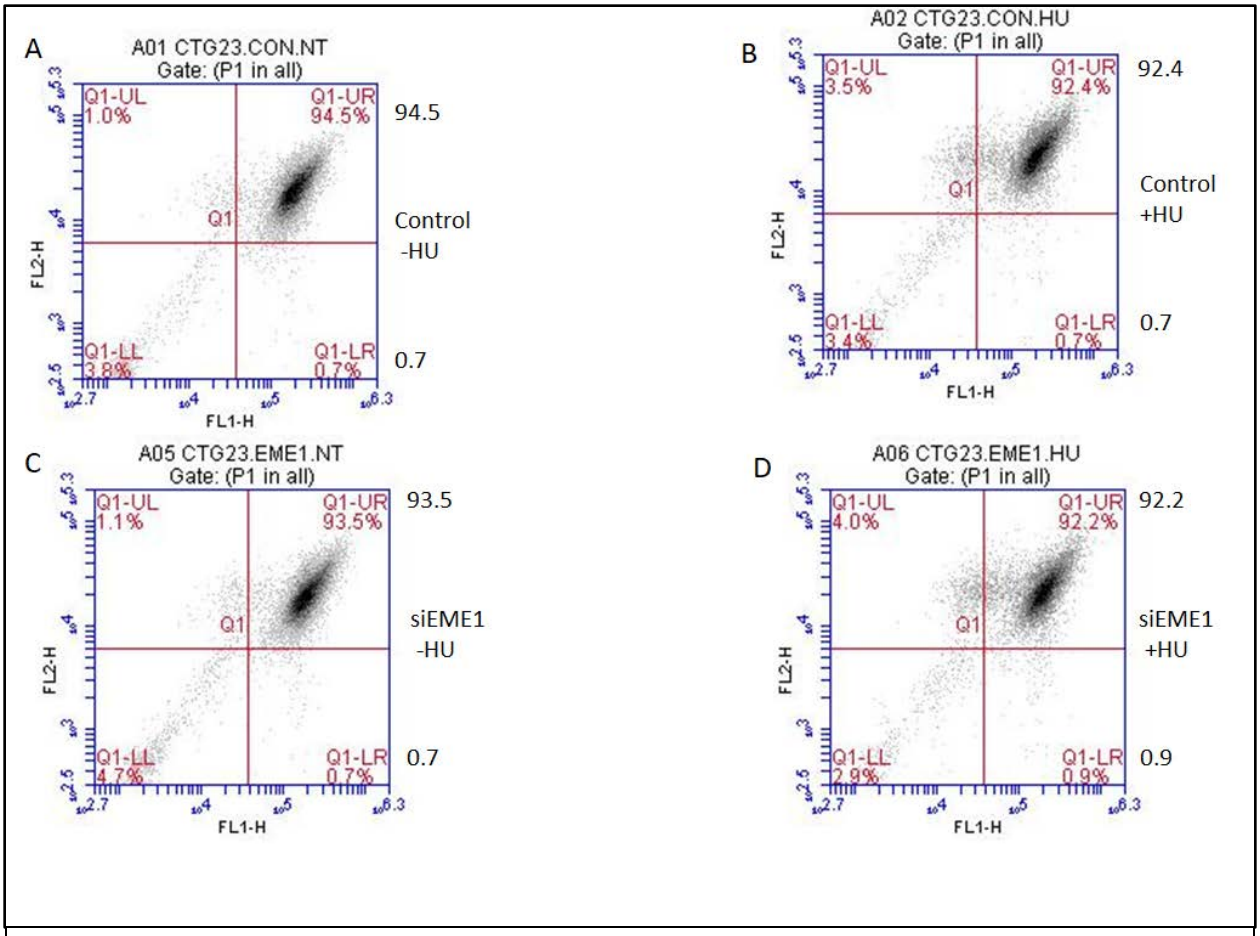



Figure 30. **siEME1 has no effect on DSBs in CTG23 cells.** (A) CTG23 cells control with no treatment of HU (B) CTG23 control cells with HU treatment. (C) CTG23 cells treated with siEME1 without HU. (D) CTG23 cells treated with siEME1 with HU.

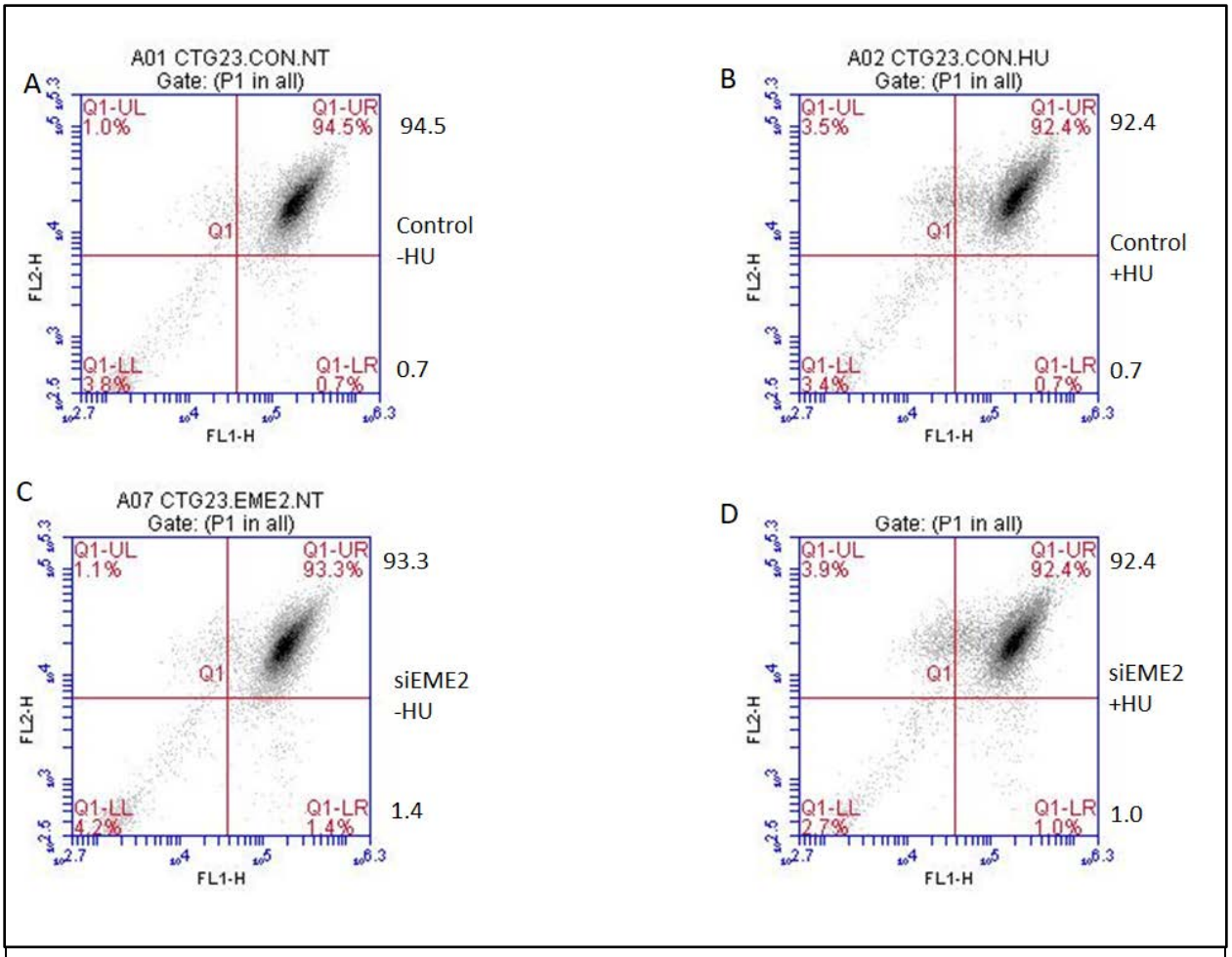



Figure 31. **siEME2 has no effect on DSBs in CTG23 cells.** (A) CTG23 cells control with no treatment of HU (B) CTG23 control cells with HU treatment. (C) CTG23 cells treated with siEME2 without HU. (D) CTG23 cells treated with siEME2 with HU.

## X. DF2 Myc $CTG_{23}$ cells show no change in ectopic microsatellite or endogenous microsatellite DMPK with spPCR

In addition to flow cytometry, it was important to look at the DNA templates of the siRNA treated CTG23 cells. After the flow cytometry profiles were taken, sample templates were purified and spPCR was performed. Primers slightly upstream and downstream of the ectopic site were used; as well as primers upstream and downstream of the endogenous DMPK locus. Additionally, internal control primers which are not associated with any microsatellites were used as an internal PCR control.

In agreement with the flow cytometry profiles shown previously, the spPCR did not show any noticeable change within the template in response to siRNA as well as hydroxyurea treatment. The contractions and expansions of the microsatellite seen within the CTG100 cell lines are not numerous enough to show in spPCR (Figure 32). Interestingly however, Figure 32 A shows new bands for all siRNAs at the ectopic site. Taken together, these results suggest that expanded CTG100 repeats are sensitive to breakage under replication stress, while the wild type length CTG repeats show expansion instability.

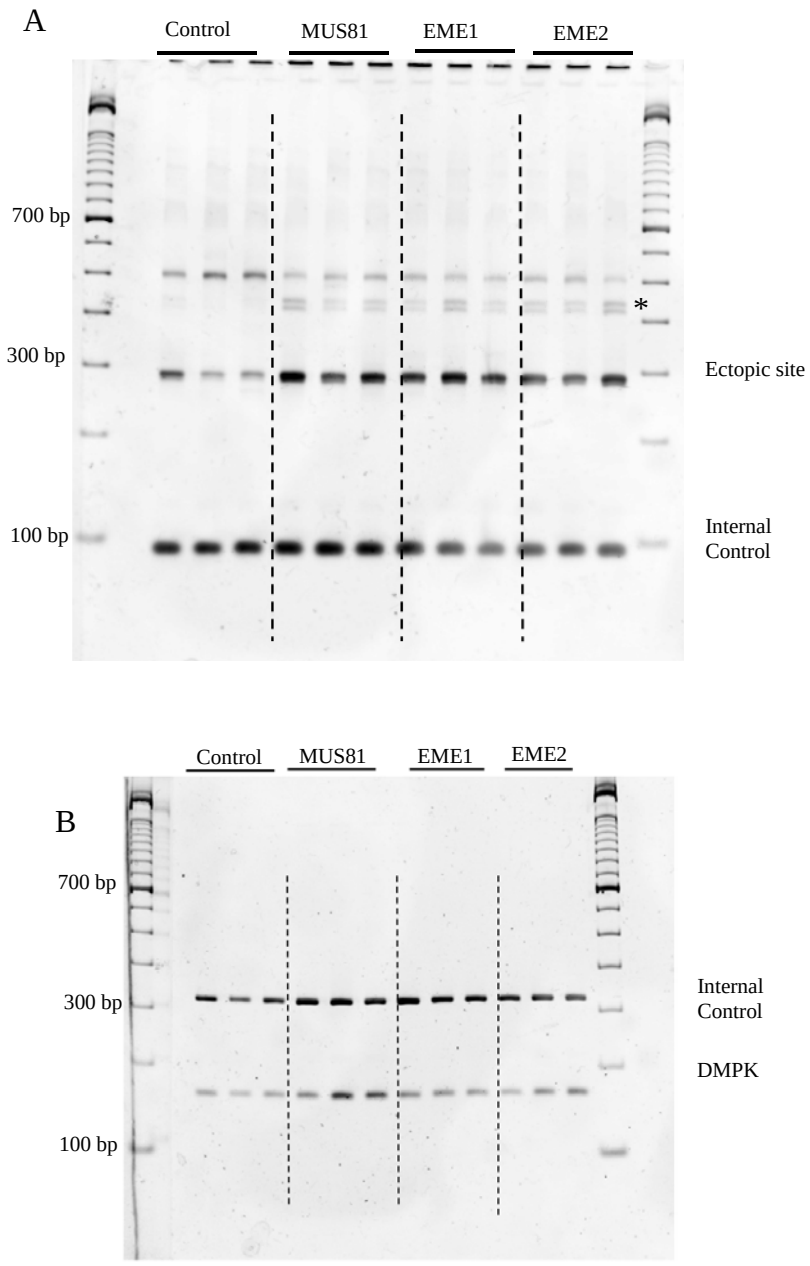



Figure 32. **siRNA treatment has no effect on spPCR of CTG23 cells.** (A) spPCR of CTG23 cells showing the ectopic site in cells treated with siMUS81, siEME1, and siEME2. (B) spPCR of CTG23 cells showing the endogenous DMPK site in cells treated with siMUS81, siEME1, and siEME2. The asterisk seen seems to indicate a tendency toward expansion instability at the ectopic site.

## Discussion

### MutSα/ MSH2-MSH6, and MutSβ/ MSH2-MSH3 loss lead to replication stress damage tolerance, and microsatellite stability

The dual fluorescence of the CTG100 and CTG23 cell lines assay the occurrence of double strand breaks that happen either spontaneously or as a result of replication stress drug treatment. It has been established in a previous study from this laboratory that the ectopic microsatellite in DF2 Myc $CTG_{100}$ cells is spontaneously unstable, and increasingly unstable under replication perturbed conditions. These assays were conducted to continue a previous study on the effects of knockdown of various proteins involved in microsatellite stability, and to see their effects on the DF2 Myc $CTG_{100}$ cell line. The work shown previously confirms the effect of HU on the CTG100 cell line compared to the CTG23 cell lines.

It has been previously established that MSH6, which is required to form MutSα is important in processing single nucleotide misalignment. In unperturbed cell cycles with knockdown of MSH6, cells showed a greater shift to double positive on the flow cytometry profile, as well as a slightly reduced response to replication stress, indicating that knockdowns of these proteins lead to stabilization within the cell line. These results agree with a study that was performed on microsatellite mutational biases due to depletion of mismatch repair proteins. It showed that MSH2 deficiency lead to single base-pair deletion bias. MSH6 deletion bias favored insertion (Shah et al 2010). These results however are at odds with a previous study by Foiry et al. (2006), which showed

that MSH6$^{-/-}$ mice do not show increased or decreased CTG/CAG microsatellite instability, whereas MSH2 and MSH3 play important roles in microsatellite expansions. Specifically, intergenerational expansion occurs in CTG repeats. These data implicates MSH3 as an illicit key in the extreme expansions observed in patients with DM1. In this study knocking down of MSH3 led to decrease in expansions of microsatellites (Foiry et al 2006).

Treatment with shMSH2 or shMSH3 (MutSβ) led to an increase in double positive cells (fewer double strand breaks) in CTG100 cells, with a greater stabilizing effect being seen with the shMSH3 assay. MSH3 has been implicated as responsible for microsatellite repeat expansions (Schmidt et al. 2016). It has also been shown that deletion of this protein favors larger deletion mutations (Shah et al. 2010), indicating that MutSβ/ MSH2-MSH3, and to a lesser extent MSH2-MSH6 (MutSα) are responsible for the instability seen within the expanded ectopic CTG microsatellite.

**Dysregulation of base excision repair pathway or crosstalk with MMR could lead to microsatellite instability**

Knockdown of Polβ showed no significant change in the percentage of double positive (yellow) cells during normal cell cycle progression. However, Polβ knockdown resulted in an increase in replication stress tolerance (fewer green cells) when treated with hydroxyurea. Previous studies suggest the reason behind these flow cytometry profiles. One study has shown that Polβ alone fails to copy trinucleotide repeats during

DNA synthesis, and will generally bypass it on a template strand which will lead to deletion. However, with stimulation by MSH2-MSH3, Polβ will copy through the repeats. Polβ stuttering leads to increased flap formation and acts as a precursor for microsatellite expansion (Lai et al. 2016) and potential fork collapse.

Another study suggests that chronic replication stress causes adaptive increases within the base excision repair pathway. This causes increased microsatellite instability suggesting that replication repair pathways that would normally help the replication process are instead causing mutation and instability (Guo et al. 2003). It is not clear whether an adaptive increase in Polβ or a lack of crosstalk between the base excision repair pathway and mismatch repair pathway could account for the increase in stability seen in the flow cytometry profiles. It is important to look at the effect of MMR and BER protein loss in depth in the future. This study helps to lay the groundwork for future experiments on this topic.

**MUS81 deficiency leads to expanded microsatellite stabilization**

The most fascinating result from all the assays performed was the MUS81 flow cytometry profile. Severe depletion of MUS81 during normal cell cycle progression led to a consistent increase in double positive cells, regardless of how many spontaneous double strand breaks the cell line was presenting prior to starting the assay. MUS81 knockdown also blunted the induction of DSBs when cells were treated with hydroxyurea. Regardless of the state the CTG100 cells were in at the beginning of the

assay, treatment with siMUS81 consistently protected the ectopic site from endogenous double strand breaks. The endogenous double strand breaks essentially stopped with the siRNA treatment, knocking down MUS81. This indicates that knockdown of MUS81 has a stabilizing effect on unperturbed cell cycle progression; it also provides protection under replication stress conditions.

A few previous studies provide evidence that agrees with this. It has been hypothesized that prolonged MUS81 activity at replication forks can lead to replication fork collapse (Rondinelli et al. 2017). This study also showed that decreased MUS81 activity at replication forks lead to fork stabilization. It has also been shown that MUS81 depletion reduces formation of double strand breaks; increases cell survival and allows S-phase progression (Forment et al. 2011). MUS81 has been shown to have a role in the rate of replication fork progression. In the absence of MUS81, DNA synthesis is slowed (Pepe et al. 2014).

The study by Fugger et al. (2013) indicates a possible signaling pathway for the toxic MUS81 function. This research presented that MUS81 and EME1 depleted cells showed increased resistance to cytotoxic stress induced by hydroxyurea. FBH1 helicase recruits MUS81 and EME1 to mediated double strand break formation in cells with prolonged replication stress, leading to activation of pro-apoptotic factor p53, restraint of RAD51 to suppress homologous recombination, and decreased double strand break formation following the treatment with hydroxyurea. Their results showed that upon siRNA mediated knockdown of MUS81 and EME1, cells treated with hydroxyurea

showed decreased double strand break formation. This model suggests that MUS81 mediated DNA double strand breaks amplify the response of checkpoint proteins leading to apoptosis in order to eliminate cells with prolonged replication stress. Figure 33 is a model based on information from Fugger et al. (2013) indicating the possible role MUS81 has in double strand breaks leading to apoptosis in cells undergoing increased amounts of replication stress.

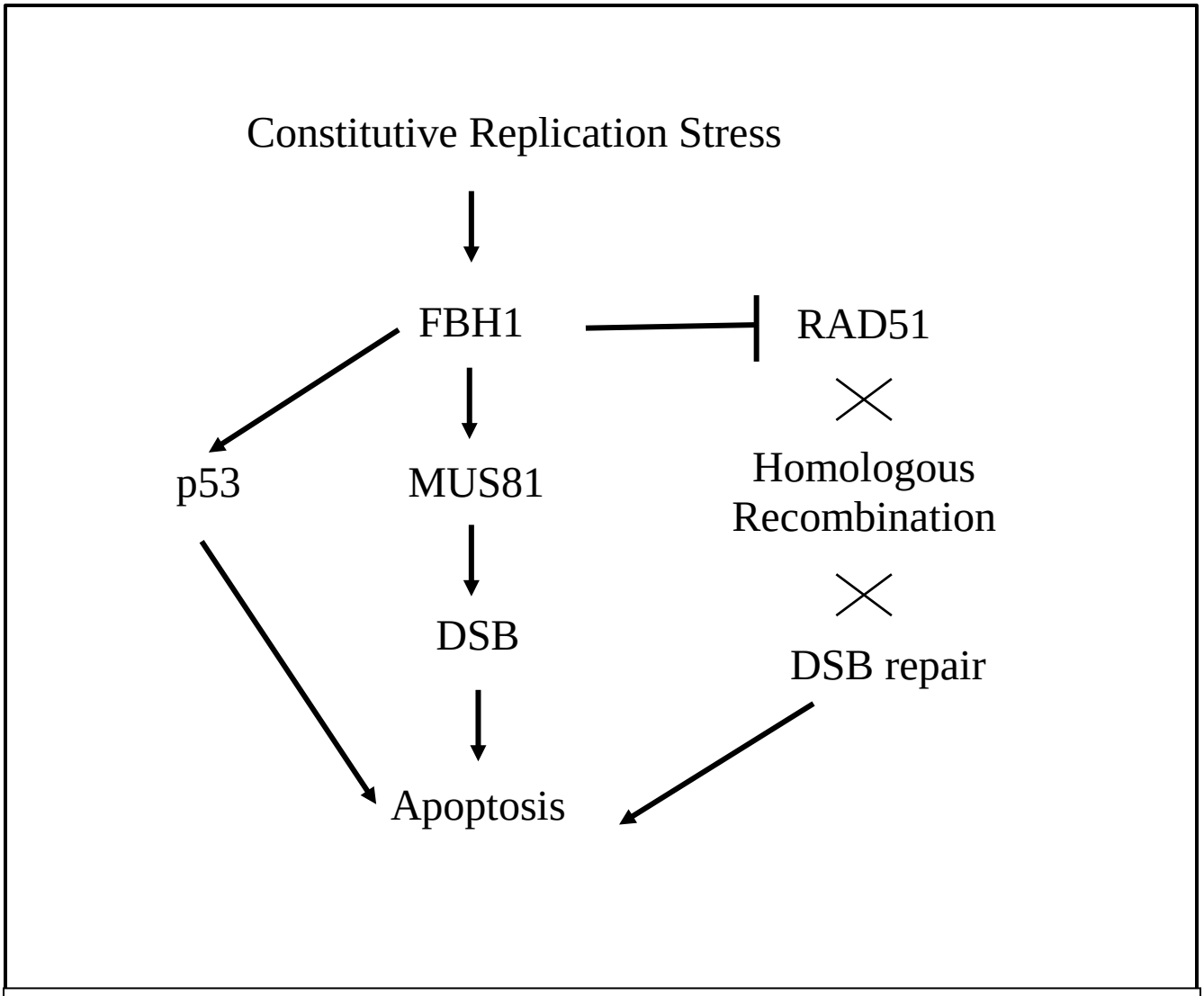



Figure 33. **Model of pathway that leads to toxic DSB through MUS81 and helicase FBH1.** Based on Fugger et. al, (2013) model indicating a toxic role of MUS81 in cells undergoing increased replication stress. Their work suggests that FBH1 helicase activates MUS81 to induce a double strand break. FBH1 also restrains RAD51, preventing double strand break repair, while simultaneously activating p53 to induce apoptosis through a toxic double strand break in cells undergoing increased, constitutive replicative stress.

**MUS81 slows cell cycle progression**

The other piece of data that seems to agree with these studies is the cell cycle progression assay with DRAQ7. At 24 hours after treatment with siMUS81, an increase in S-phase cells is seen, but at 48 hours the G1 peak has decreased, seeming to indicate a slowing of cell cycle progression. This in combination with the flow cytometry profile seems to suggest that in the DF2 Myc $CTG_{100}$ cell lines knockdown of MUS81 leads to slowed cell cycle progression that has a stabilizing effect on the expanded microsatellite during unperturbed growth.

MUS81 null cancer cells during unperturbed growth show a higher rate of early replication fork initiation and slow DNA replication. In the study by Fu et al. (2015), their observations indicated a couple possibilities for this. One possibility is that loss of MUS81 results in slow replication, which consequently increases the initiation of early replication forks. The second possibility is that knockdown of MUS81 facilitates replication fork initiation, slowing replication indirectly (Fu et al. 2015). The stabilizing effect of MUS81 on expanded microsatellites seems to indicate the toxic effect MUS81 has on these cells. This could be because knockdown of MUS81 has a slowing rather than stalling effect on the replication fork, and/or because in the CTG100 cells might have MUS81 that has become over active in areas of stalling creating a toxic effect on these cells.

**EME2 rather than EME1 plays a more prominent role in replication fork stabilization**

Another unexpected outcome was the result of the siRNA treated EME1 and EME2 DF2 Myc $CTG_{100}$ cells. Knockdown with EME1 had a mild stabilizing effect on the cells compared to the control. The double positive cells would show slight increase compared to the control, and a slight decrease in eGFP cells when treated with hydroxyurea. More interestingly however, was that knockdown of EME2 drastically increased replication instability. The CTG100 cell line consistently had a drastic increase in eGFP cells compared to control, both with and without hydroxyurea treatment. Thus, EME2 may have a major role in stabilizing the trinucleotide region and microsatellite stability, consistent with EME2 activity during S phase. Alternatively, decreased levels of EME2 may increase the level of MUS81-EME1 complexes at an inappropriate time during S phase.

This increase in instability of the microsatellite regions was confirmed with the spPCR data. Only cells treated with siEME2 showed expansion and contraction, not only in the expanded exogenous microsatellite, but also at the endogenous DMPK locus. These results only appeared in a small percentage of the cells, but it indicates a larger role for MUS81-EME2 in the stability of both long and short microsatellites.

**EME1 and EME2 siRNA knockdown caused S phase arrest**

Almost at odds with the flow cytometry profile data was the cell cycle progression assay using DRAQ7. This indicated that knockdown of EME1 and EME2 both caused S-phase cell cycle arrest in the DF2 cell line; however, EME2 had much more drastic consequences. These results seem to agree with another study that suggests that MUS81-EME2 complexes are more active and more versatile than the MUS81-EME1 complex. This study showed that MUS81-EME2 was able to process Holliday junctions as well as nicked and gapped duplex DNA. It also showed that EME2 is involved in the repair of both stalled and regressed replication forks (Amangyeld et al 2014). This seems to indicate an active, significant role for EME2 in S phase.

The knockdown of EME2 enhances endogenous double strand breaks, and is comparable in the amount of cells with DSBs compared to HU treatment. Hydroxyurea is known to cause replication stress during S phase of the cell cycle. The mechanism of hydroxyurea replication stress indicates that HU causes defects in the timing of origin activation and fork progression. The replication forks are generated at early firing origins, and become hindered by the low levels of dNTPs (Wang, Vasquez 2009).

It is possible that EME2 and HU affect the same pathway of replication, acting and affecting the cell cycle during S phase. Hydroxyurea acts by disrupting the DNA replication process by inhibiting ribonucleotide reductase, which is an enzyme required to convert ribonucleoside diphosphates into deoxyribonucleoside diphosphates. This inhibits cells from leaving the S phase of the cell cycle. In addition to its effect on stalled

replication forks, MUS81-EME2 dependent DNA damage occurs through activation of the ATM pathway when CHK1 is downregulated. The inability of CHK1 to activate ribonucleotide reductase seems to limit dNTPS available for replication which results in slower replication fork progression (Pai and Kearsey 2017, Techer et al. 2016).

In the study by Amangyeld et al. (2013) it was indicated that MUS81-EME2 cleaves Holliday junctions more efficiently compared to MUS81-EME1. It also cleaves a variety of other DNA structures that arise through replication. The flow cytometry results seem to agree with this. Knockdown of EME2 in the CTG100 cells, which are prone to formation of non-B DNA structures, seems to have greatly hindered the cells ability to replicate properly. The added data that EME2 knockdown is comparable to HU treatment in terms of increase in DSBs further indicate the occurrence of instability in S phase implicating the importance of EME2 in maintaining stability in these cells.

As mentioned previously the endonuclease MUS81 is needed for EME1 and EME2 activation. It is curious that MUS81 knockdown gives different results than that of siEME1 and siEME2. This could happen because of the different stages of the cell cycle these proteins are active in. MUS81-EME2 is active during S phase of the cell cycle while MUS81-EME1 is active during G2/M stages of the cell cycle. The knockdown of EME1 seemed to have created a slight protection of the microsatellite (indicated by slight decrease in double strand breaks occurring with decrease in number of green cells and increased yellow cells). This seems to suggest that the toxic effect of MUS81-EME1

could be occurring during G2/M portion of the cell cycle rather than during S phase during active DNA replication.

MUS81 depletion and overexpression has been linked to the rate of replication fork progression (Mayle et al. 2015). In addition to affecting the rate of replication fork progression MUS81 has been implicated in activating the intra-S-phase checkpoint and the G2/M checkpoint during cell cycle. With MUS81 and EME1 knockdowns displaying increased stability within the genome, it suggests that there could be a type of functional redundancy between the MUS81-EME1 complex and other DNA damage response proteins during the G2/M portions of the cell cycle (Hiyama et al. 2006). The increased instability seen in the CTG100 cells when treated with siEME2 indicates that EME2 is vital for maintenance of genetic stability at the expanded microsatellite.

Expansions and contractions of trinucleotide repeats have been studied extensively. The CTG100 cell lines show that the expanded trinucleotide repeat can also be considered a fragile site due to non-canonical DNA structure formation that can lead to chromosome breaks during unperturbed replication, and during replication stress. The CTG100 cell line display increased spontaneous DNA breaks, and recombination which is detectable by a shift in fluorescence. Within the CTG100 cell line, the expanded microsatellite sequence was placed between the fluorescent color markers. In the event of an unrepaired break at the microsatellite sequence, loss of color marker protein will allow for a shift in fluorescence to be observed. The previous shown results suggest that DNA secondary structures can inhibit DNA recombination after a DSB near a microsatellite

motif. The CTG100 cell line can be used to detect DNA repair patterns. It is also a novel system that allows us to test the effects non-B DNA structures have on producing DNA double strand breaks, or influencing the repair pathway that is utilized within human cells. The DF2 Myc $CTG_{100}$ cell line is an ideal system for detection of DNA repair patterns, both homologous and non-homologous. As well as how DNA microsatellites and non-canonical DNA structures effect DNA repair, and double strand breaks.

In conclusion, the loss of crossover junction endonucleases has a varied effect on DNA stability. Depletion of MUS81 leads to increased stability within the cell, possibly because its knockdown causes slowed cell cycle progression. The knockdown might also diminish a toxic effect which could be caused by over activity, or due to a downstream signal which leads to cellular apoptosis due to prolonged replication stress. Loss of EME2 greatly reduces replication stability and the cell's ability to handle DNA replication stress and damage. This shows that a repair mechanism could lead to microsatellite instability, while also being important for cellular survival.

**Break Induced Replication**

An additional aspect of this study we would like to explore is the effect of these protein knockdowns in a new cell line that is being developed that demonstrates break induced replication. Break induced replication (BIR) can account for extremely large microsatellite repeat expansions and gross chromosomal rearrangements that occur within a single replication event (Kramara et al. 2017). Break induced replication is a

type of homologous recombination (HR) that is seen in all organisms from prokaryotes to eukaryotes. It is a distinct mechanism that can lead to large scale repeat expansion in a single step. BIR is initiated by strand invasion which is instigated by Rad51 to form a displacement loop (D-loop) (Kim et al. 2016). This establishes a replication fork consisting of both leading and lagging strand synthesis. This process can continue to the end of replication or even to the end of the chromosome arm, and leads to conservative replication (Fig. 33). Initial extension from a double strand end has been shown to require a subunit of Polδ, Pol32 (Hastings et al. 2009). Collapsed replication fork repair often uses the error-prone mediated pathway using Pol32 dependent synthesis. However this synthesis is usually limited to a few kilobases from the break by MUS81 endonuclease and the converging fork. MUS81 suppresses template switches between both homologous sequence and human Alu repetitive elements (Mayle et al. 2015).

The subsequent strand invasion and delayed lagging strand lead to conservative replication. This process is highly mutagenic, considered to be 1000x greater than semiconservative replication (Malkova and Ira 2013). It has been proposed that BIR repair of damaged replication forks during DSR events is responsible for the high frequency of genomic duplications, as well as the high rate of mutations caused by addition of DSB and accumulation of genome instability seen in human cancers (Costantino et al. 2014, Sakofsky et al. 2012).

In budding yeast, it has been shown that BIR can be considered one of the mechanisms that leads to expansion of (CAG)•(CTG) microsatellites. Kim et al.

demonstrated that the non-B DNA structures, which lead to chromosomal instability and the misalignment of the DNA strands during strand invasion, leads to expansion by BIR. They also demonstrated, through genetic knockouts, proteins that are specific for large scale expansions are involved in this process, whereas proteins important for small scale expansions are not involved. In this study, protein knockouts involved in homologous recombination including RAD51, RAD52, MUS81, YEN1, MRE11 as well as proteins involved in break induced replication including POL32, and PIF1 decreased (CAG)•(CTG) expansions. The proteins responsible for small scale expansions including MSH2, MSH3, MSH6, did not affect the expanded microsatellite (Malkova and Ira 2013). In budding yeast it has been shown that break induced replication repairs collapsed replication forks in cells with oncogene induced DNA repair stress through the use of Rad52 and Pol32 (Sotirious et al. 2016, Costantino et al. 2014). For a future project, after this cell line has been established it would be interesting to see the effects of these knock downs on the cell line and see if they agree with previous studies in other experimental systems.

Figure 34 shows the construct of this cell line called DF2 eGFP-TK fusion. An enzymatic digest has been performed on the eGFP-TK fusion plasmid to confirm the structure of the construct (Figure 35). This cell line will be very similar to the DF2 Myc $CTG_{100}$ cell line used previously. The only difference is that the eGFP gene will not have a stop codon at its end. This will allow replication to continue into the TK region of the 406 HeLa cells, and form an eGFP-TK fusion protein. This fusion will make the cells

sensitive to ganciclovir treatment; however, when BIR mutagenesis occurs (Figure 36) the cells will become ganciclovir resistant allowing for selection of the cells that have undergone break induced replication.

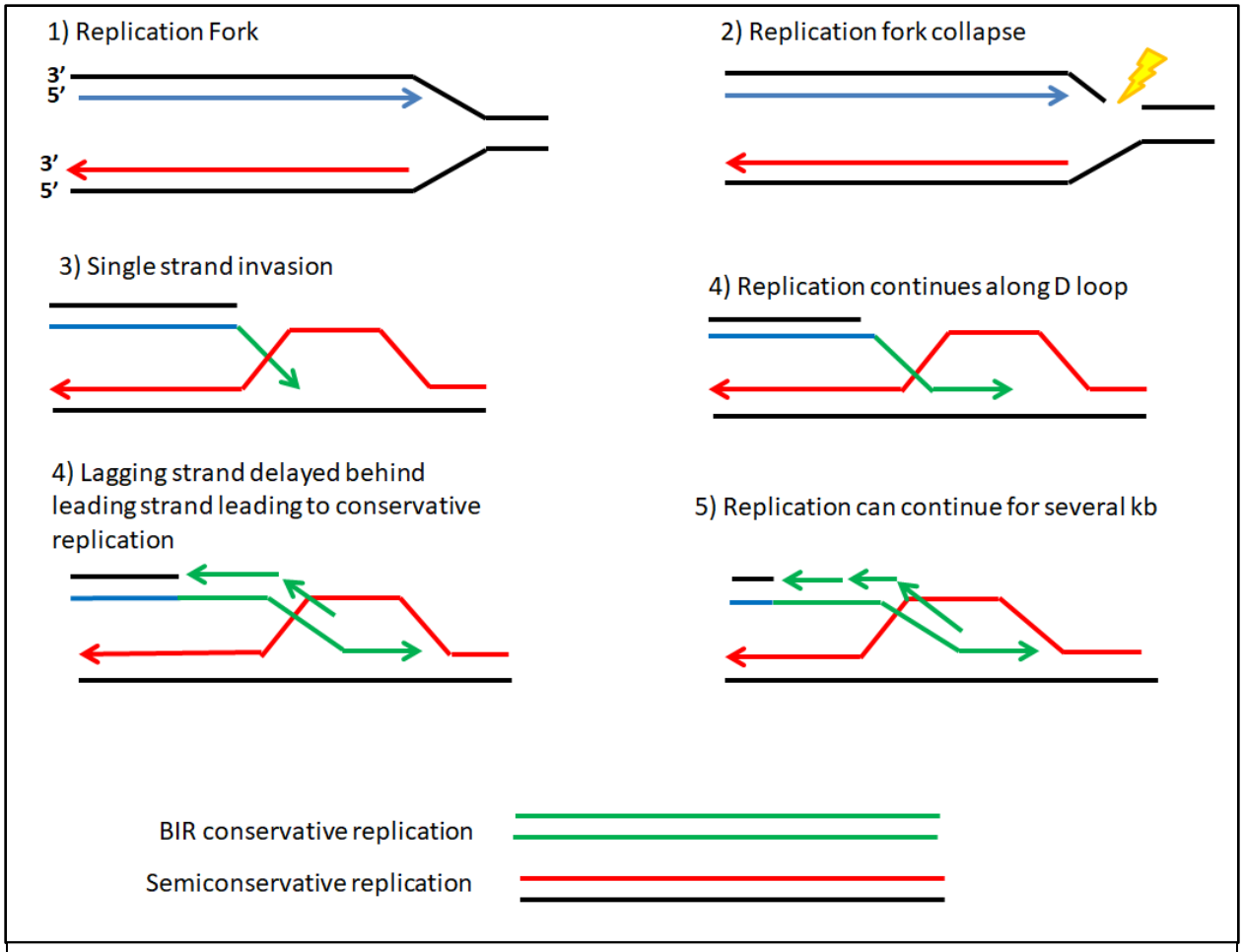



Figure 34. **Model representing how Break Induced Replication occurs.** Model showing how BIR leads to conservative replication and is considered to be highly mutagenic form of replication.

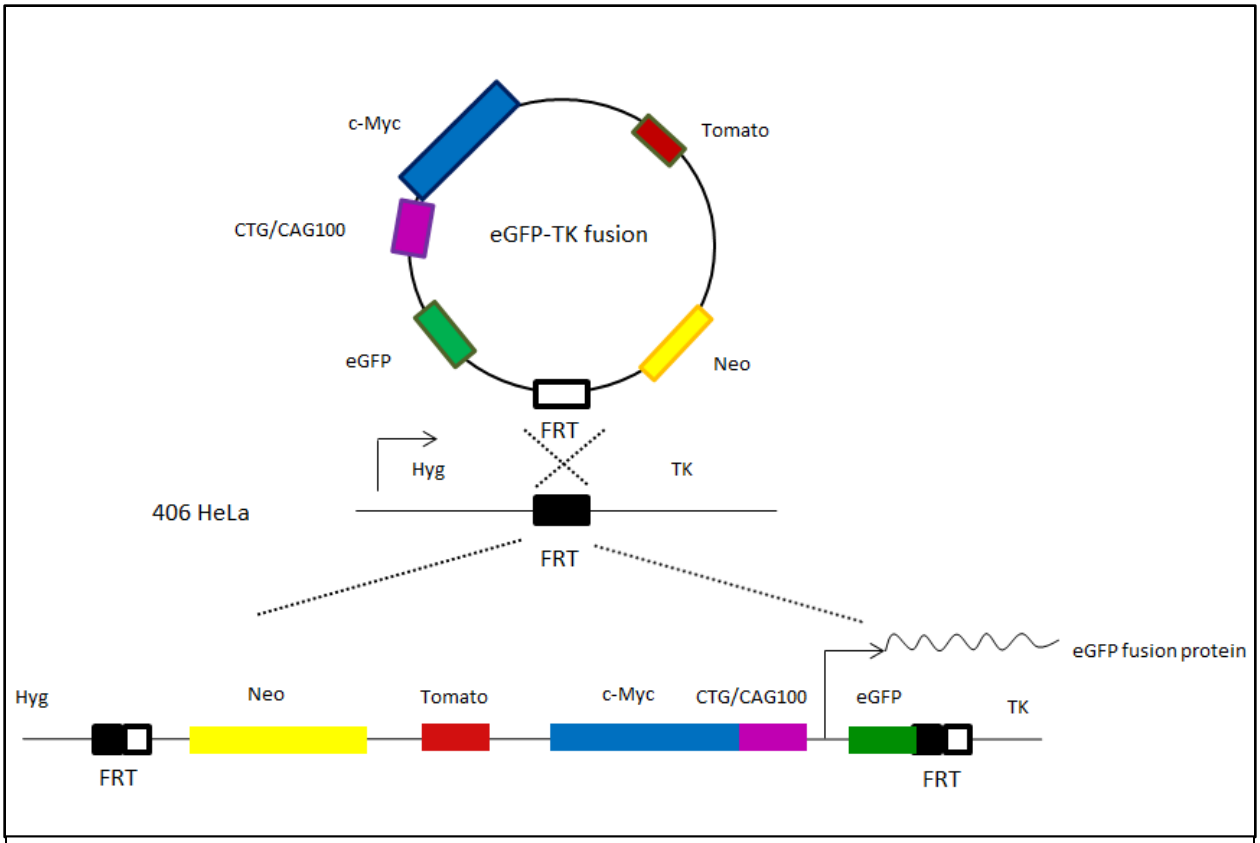



Figure 35. **Model of DF2 eGFP-TK fusion cell line.** Model of eGFP-TK plasmid and integration into HeLa 406 cells leading to the new cell line that forms a eGFP-TK fusion protein.

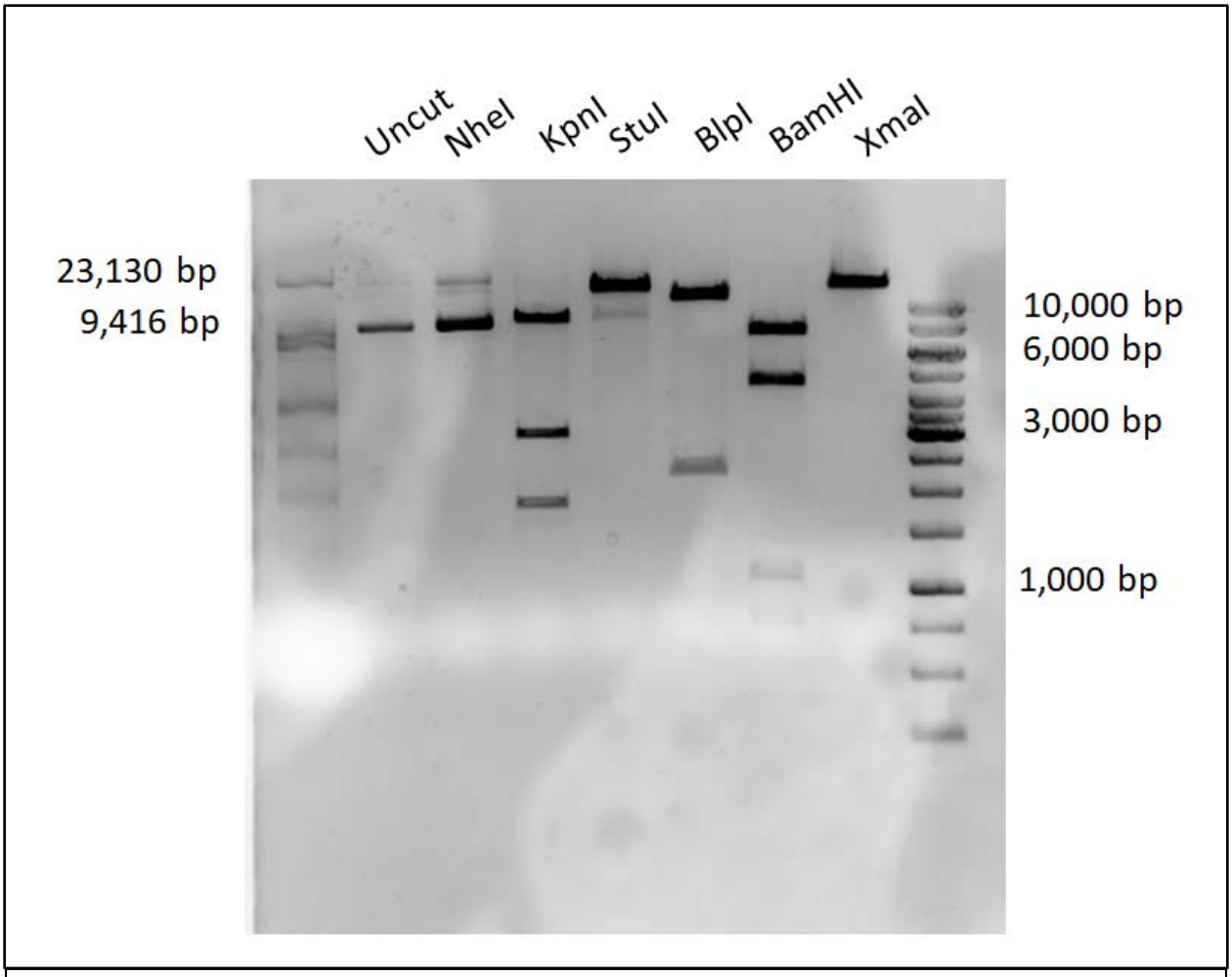



Figure 36. **Enzyme digest of eGFP-TK fusion plasmid.** NheI- 0 cutter enzyme to check that plasmid is supercoiled. KpnI- 4 cutter enzyme, cuts close to eGFP. StuI 2 cutter enzyme that cuts within dTomato. BlpI- 2 cutter enzyme that cuts close to dTomato. BamHI- 4 cutter enzyme that cuts close to dTomato and eGFP on either side of the color markers. XmnI- single cutter enzyme to check plasmid length.

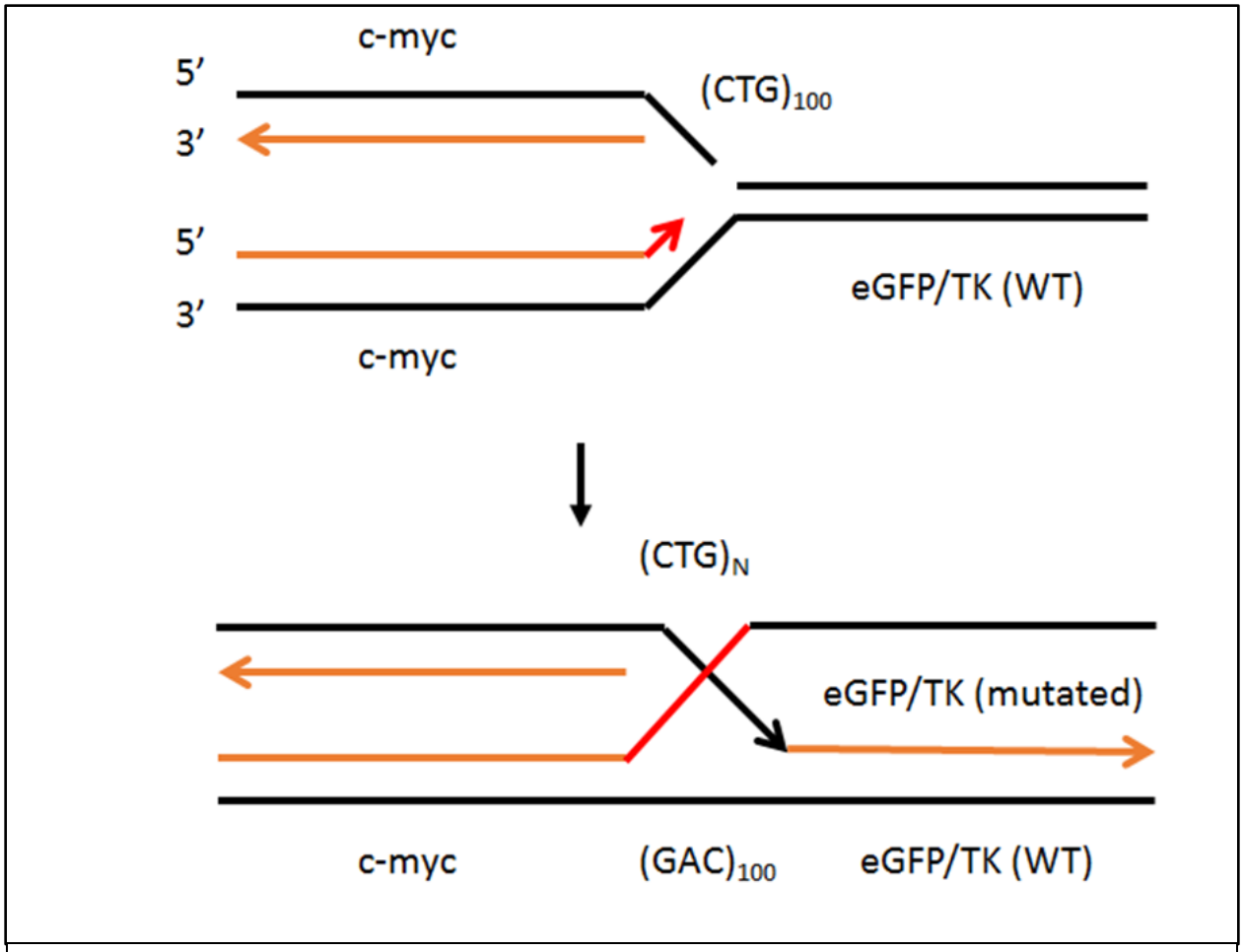



Figure 37. **Model of how the eGFP-TK cell line will lead to BIR.**

**Future Directions**

In the future, we would like to study the effect of knocking down various mismatch repair proteins and base excision repair proteins and their effect on expanded microsatellites in more detail. These proteins have an effect on the cells containing expanded trinucleotide repeats, and seeing these effects through spPCR, and using next generation sequencing to map the mutations would be a worthwhile endeavor.

It would also be fascinating to explore the effect of other DNA replication stressing drugs in combination with the siRNA treatment of the CTG100 cells. Especially interstrand cross-linking agents with a MUS81 siRNA knockdown. It would be interesting to see if siRNA mediated knockdown of MUS81 increases the CTG100 and CTG23 cells sensitivity to interstrand cross-linking instability as has been shown in other studies with various cancer cells lines.

To further expand on this project and the proteins examined would be to study the DF2 Myc $CTG_{100}$ cell line and the DF2 Myc $CTG_{23}$ cell lines basal protein level for comparison. As mentioned previously, and in many other studies, particularly studies involving mismatch repair proteins and base excision repair proteins, these pathways become dysregulated which can lead to microsatellite instability. It would be interesting to examine if one of the proteins used in this project changes levels which could explain why the cell line becomes progressively more unstable, or if the toxic MUS81 effect seen is part of a signaling cascade initiated by FBH1 to lead to apoptosis in cells with constitutive replication stress. A similar assay would be to measure DNA damage

response proteins levels in response to siRNA treatment, such as γH2AX, pCHK1, and pCHK2. It would also be interesting to look at other microsatellite besides DMPK and the ectopic CTG site, as well as explore site-specific protein mutants involving MUS81, EME1, and EME2. Further exploration into this project could give key implications in the study of oncogenic genomic instability.

**References**


1. Amangyeld, T., Shin, Y., Lee, M., Kwon, B., & Seo, Y. (2014). Human MUS81-EME2 can cleave a variety of DNA structures including intact Holliday junction and nicked duplex, *42*(9), 5846–5862. https://doi.org/10.1093/nar/gku237
2. Axford, M. M., Wang, Y. H., Nakamori, M., Zannis-Hadjopoulos, M., Thornton, C. A., & Pearson, C. E. (2013). Detection of Slipped-DNAs at the Trinucleotide Repeats of the Myotonic Dystrophy Type I Disease Locus in Patient Tissues. *PLoS Genetics*, *9*(12). https://doi.org/10.1371/journal.pgen.1003866
3. Barakat, K., Gajewski, M., & Tuszynski, J. (2012). DNA polymerase beta (pol β) inhibitors: A comprehensive overview. *Drug Discovery Today*, *17*(1), 913–920.
4. Barthelemy, J., Leffak, M., & Hanenberg, H. (2016). FANCJ is essential to maintain microsatellite structure genome-wide during replication stress. *Nucleic Acids Research*\, *44*(14), 6803–6816.
5. Cabelof, D., Guo, Z., Raffoul, J., Sobol, R., Wilson, S., Richardson, A., & Heydari, A. (2003). Base excision repair deficiency caused by polymerase beta haploinsufficiency: accelerated DNA damage and increased mutational response to carcinogens. *Cancer Res.*, *63*(18), 5799–5807.
6. Carignon, S., Lopez, B. S., Brison, O., & Debatisse, M. (2016). Signaling from Mus81-Eme2-Dependent DNA Damage Elicited by Chk1 Deficiency Modulates Replication Fork Speed and Origin Usage Article Signaling from Mus81-Eme2-

Dependent DNA Damage Elicited by Chk1 Deficiency Modulates Replication Fork Speed and Origin Usage, 1114–1127. https://doi.org/10.1016/j.celrep.2015.12.093

7. Chatterjee N, Lin Y, Santillan BA, Yotnda P, Wilson JH: Environmental stress induces trinucleotide repeat mutagenesis in human cells. Proc Natl Acad Sci US A 2015, 112:3764e3769

8. Chen, J. M., Chuzhanova, N., Stenson, P. D., Férec, C., & Cooper, D. N. (2005). Complex gene rearrangements caused by serial replication slippage. *Human Mutation*, *26*(2), 125–134. https://doi.org/10.1002/humu.20202

9. Ciccia, A., Constantinou, A., & West, S. C. (2003). Identification and Characterization of the Human Mus81-Eme1 Endonuclease *, *278*(27), 25172–25178. https://doi.org/10.1074/jbc.M302882200

10. Costantino, L., Sotiriou, S. K., Rantala, J. K., Magin, S., Mladenov, E., Helleday, T., Halazonetis, T. D. (2014). Break-induced replication repair of damaged forks induces genomic duplications in human cells. *Science (New York, N.Y.)*, *343*(6166), 88–91. https://doi.org/10.1126/science.1243211

11. Cummings, C.J. and Zoghbi, H.Y. (2000) Fourteen and counting: unravelling trinucleotide repeat diseases. *Hum.Mol.Genet.*, 9 , 909-916.

12. Dendouga, N., Gao, H., Moechars, D., Janicot, M., Vialard, J., & Mcgowan, C. H. (2005). Disruption of Murine Mus81 Increases Genomic Instability and DNA Damage Sensitivity but Does Not Promote Tumorigenesis, *25*(17), 7569–7579. https://doi.org/10.1128/MCB.25.17.7569

13. Edge, L. (2017). Review Endogenous DNA Damage as a Source of Genomic Instability in Cancer, 644–656. https://doi.org/10.1016/j.cell.2017.01.002

14. Ellard, S., Bulman, M., Frayling, T., Shepherd, M., & Hattersly, A. (2000). Proposed Mechanism for a Novel Insertion/Deletion Frameshift Mutation (I414G415ATCG→→CCA) in the Hepatocyte Nuclear Factor 1 Alpha (HNF-1αα) Gene Which Causes Maturity-Onset Diabetes of the Young(MODY). *Human Mutation: Mutation in Brief, 360*. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/1098-1004(200009)16:3%3C273::AID-HUMU18%3E3.0.CO;2-Z/epdf

15. Entrez Gene: POLB polymerase (DNA directed), beta https://www.ncbi.nlm.nih.gov/gene?Db=gene&Cmd=ShowDetailView&TermToSearcT=5423

16. Fadden, A. J., Schalbetter, S., Bowles, M., Harris, R., Lally, J., Carr, A. M., & Mcdonald, N. Q. (2013). A winged helix domain in human MUS81 binds DNA and modulates the endonuclease activity of MUS81 complexes, *41*(21), 9741–9752. https://doi.org/10.1093/nar/gkt760

17. Foiry, L., Dong, L., Savouret, C., Hubert, L., Riele, H.T., Junien, C., and Gourdon, G. (2006) Msh3 is a limiting factor in the formation of intergenerational CTG expansions in DM1 transgenic mice. Hum. Genet. *119*, 520-526.)

18. Forment JV, Blasius M, Guerini I, Jackson SP (2011) Structure-Specific DNA Endonuclease Mus81/Eme1 Generates DNA Damage Caused by Chk1 Inactivation. PLoS ONE 6(8): e23517. doi:10.1371/journal.pone.0023517

19. Fu, H., Martin, M. M., Regairaz, M., Huang, L., You, Y., Lin, C., … Aladjem, M. I. (2015). The DNA repair endonuclease MUS81 facilitates fast DNA replication in the absence of exogenous damage. *Nature Communications*, *6*, 1–14. https://doi.org/10.1038/ncomms7746

20. Fugger, Kasper, Wai Kit Chu, Peter Haahr, Arne Nedergaard Kousholt, Halfdan Beck, Miranda J Payne, Katsuhiro Hanada, Ian D. Hickson and Claus S Sørensen. "FBH1 co-operates with MUS81 in inducing DNA double-strand breaks and cell death following replication stress." *Nature communications* 4 (2013): 1423.

21. Gacy AM, Goellner G, Juranic N, Macura S, McMurray CT. Trinucleotide repeats that expand in human disease form hairpin structures in vitro. Cell. 1995;81:533–540. doi: 10.1016/0092-8674(95)90074-8.

22. Gadgil, R. Barthelemy, J. Lewis, T. Leffak, M. (2016). Replication stalling and DNA microsatellite instability. *Biophysical Chemistry*.

23. Ghamrasni, S. El, Pamidi, A., Halaby, M. J., Bohgaki, M., Cardoso, R., Li, L., … Hakem, R. (2011). Inactivation of Chk2 and Mus81 Leads to Impaired Lymphocytes Development , Reduced Genomic Instability , and Suppression of Cancer, *7*(5). https://doi.org/10.1371/journal.pgen.1001385

24. Goula,A.V. and Merienne,K. (2013) Abnormal base excision repair at trinucleotide repeats associated with diseases: a tissue-selective mechanism. Genes, 4, 375–387.

25. Gupta, S., Gellert, M., & Yang, W. (2012). Mechanism of mismatch recognition revealed by human MutSβ bound to unpaired DNA loops. *Nature Structural and Molecular Biology*, *19*, 72–78.

26. Haiqing, F., Martin, M., & Aladjem, M. (2015). The DNA repair endonuclease Mus81 facilitates fast DNA replication in the absence of exogenous damage. *Nature Communications*, *6*, 6746.

27. Hastings, P., Lupski, J. R., Rosenberg, S. M., & Ira, G. (2009). Mechanisms of change in gene copy number. *Nature Reviews. Genetics*, *10*(8), 551–564. https://doi.org/10.1038/nrg2593

28. Hiyama, T., Katsura, M., Yoshihara, T., Ishida, M., Kinomura, A., Tonda, T., … Miyagawa, K. (2017). Haploinsufficiency of the Mus81 – Eme1 endonuclease activates the intra-S-phase and G 2 / M checkpoints and promotes rereplication in human cells, *34*(3), 880–892. https://doi.org/10.1093/nar/gkj495

29. Hiyama, T., Katsura, M., Yoshihara, T., Ishida, M., Kinomura, A., Tonda, T., … Miyagawa, K. (2006). Haploinsufficiency of the MUS81-Eme1 endonuclease activates the intra-S-phase and G2/M checkpoints and promotes rereplication in human cells. *Nucleic Acids Research*, *34*(3), 880–892.

30. Holloway, J. K., Booth, J., Edelmann, W., Mcgowan, C. H., & Cohen, P. E. (2008). MUS81 Generates a Subset of MLH1-MLH3 – Independent Crossovers in Mammalian Meiosis, *4*(9). https://doi.org/10.1371/journal.pgen.1000186

31. Idriss, H., Al-Assar, O., & Wilson, S. (2002). DNA polymerase Beta. *The International Journal of Biochemistry and Cell Biology*, *34*(4), 321–324.

32. Jane C Kim, Sergei M Mirkin. (2013). The balancing act of DNA repeat expansions. *Current Opinion in Genetics and Development*, *23*(3), 280–288. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3703482/

33. Jinzhen Guo, Luping Chen, Guo-Min Li, DNA mismatch repair in trinucleotide repeat instability, SCIENCE CHINA Life Sciences, 60, 1087 (2017)).

34. Kim, J. C., Harris, S. T., Dinter, T., Shah, K. A., & Mirkin, S. M. (2016). The role of break-induced replication in large-scale expansions of (CAG)n/(CTG)n repeats. *Nature Structural & Molecular Biology*, (December). https://doi.org/10.1038/nsmb.3334

35. Kramara, J., Osia, B., and Malkova, A. (2017). Break-induced replication: an unhealthy choice for stress relief? Nat Struct Mol Biol *24*, 11-12.

36. Lai, Y., Budworth, H., Beaver, J. M., Chan, N. L. S., Zhang, Z., & Mcmurray, C. T. (2016). during base excision repair. *Nature Communications*, *7*, 1–15. https://doi.org/10.1038/ncomms12465

37. Lee, D.-Y., & McMurray, C. (2014). Trinucleotide expansion in disease: why is there a length threshold? *Current Opinion in Genetics and Development*, 130–140. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4252851/

38. Lewis, S. M. (2008). DNA Breaks at In Vivo-Generated Cruciform Structures in S . cerevisiae, 800–812. https://doi.org/10.1016/j.molcel.2008.08.025

39. Liu, G., & Leffak, M. (2012). Instability of (CTG)n•(CAG)n trinucleotide repeats and DNA synthesis. Cell & Bioscience, 2(1), 7. https://doi.org/10.1186/2045-3701-2-7

40. Liu, G., Chen, X., Bissler, J. J., Sinden, R. R., & Leffak, M. (2010). Replication-dependent instability at (CTG)•(CAG) repeat hairpins in human cells. *Nature Chemical Biology*, *6*(9), 652–659. https://doi.org/10.1038/nchembio.416

41. Liu, X., Krawczyk, E., Suprynowicz, F., Palechor-Ceron, N., Yuan, H., Dakic, A., Schlegel, R. (2017). Conditional reprogramming and long-term expansion of normal and tumor cells from human biospecimens. *Springer Nature*, *12*(2).

42. Malkova, A., & Ira, G. (2013). Break-induced replication: Functions and molecular mechanism. *Current Opinion in Genetics and Development*. https://doi.org/10.1016/j.gde.2013.05.007

43. Mallon, M. A., Zhang, W., Freimuth, R. R., Marsh, S., Watson, M. A., Goodfellow, P. J., & Mcleod, H. L. (2006). Human Cancer Biology DNA Repair Pathway Profiling and Microsatellite Instability in Colorectal Cancer, *12*(17), 5104–5112. https://doi.org/10.1158/1078-0432.CCR-06-0547

44. Manuscript, A., Repair, D. M., & Bias, M. M. (2011). NIH Public Access, *70*(2), 431–435. https://doi.org/10.1158/0008-5472.CAN-09-3049.Defective

45. Mayle, R., Campbell, I., Beck, C., Yu, Y., Wilson, M., Shaw, C., … Ira, G. (2015). MUS81 and converging forks limit the mutagenicity of replication fork breakage. *DNA Repair*, *349*(6249).

46. Mellon I, Rajpal DK, Koi M, Boland CR, Champe GN (April 1996). "Transcription-coupled repair deficiency and mutations in human mismatch repair genes". Science. 272 (5261): 557–60. doi:10.1126/science.272.5261.557. PMID 8614807.

47. Mirkin, S. M. (2006). DNA structures, repeat expansions and human hereditary disorders. *Current Opinion in Genetics and Development*, *16*(3), 351–358.

48. Mirkin, S. M. (2007). Expandable DNA repeats and human disease. *Nature*, *447*(7147), 932–940. https://doi.org/nature05977 [pii]\r10.1038/nature05977

49. Mott, C., & Symington, L. S. (2011). RAD51-independent inverted-repeat recombination by a strand-annealing mechanism. *DNA Repair*, *10*(4), 408–415. https://doi.org/10.1016/j.dnarep.2011.01.007

50. Oke, A., Anderson, C. M., Yam, P., & Fung, J. C. (2014). Controlling Meiotic Recombinational Repair – Specifying the Roles of ZMMs , Sgs1 and Mus81 / Mms4 in Crossover Formation, *10*(10). https://doi.org/10.1371/journal.pgen.1004690

51. Pai, Chen-Chun, and Stephen E. Kearsey. “A Critical Balance: dNTPs and the Maintenance of Genome Stability.” Ed. Eishi Noguchi. *Genes* 8.2 (2017): 57. *PMC*. Web. 19 Feb. 2018.

52. Pelletier, R., Farrell, B. T., Miret, J. J., and Lahue, R. S. (2005) Mechanistic features of CAG*CTG repeat contractions in cultured cells revealed by a novel genetic assay. Nucleic Acids Res. 33, 5667–567.

53. Pelletier, R., Krasilnikova, M., Samadashwily, G., Lahue, R., & Mirkin, S. (2003). Replication and expansion of trinucleotide repeats in yeast. *Molecular and Cellular Biology*, *23*(4), 1349–1357.

54. Pepe, A., & West, S. C. (2013). Substrate specificity of the MUS81-EME2 structure selective endonuclease, 1–13. https://doi.org/10.1093/nar/gkt1333

55. Pepe, A., & West, S. C. (2014). Report MUS81-EME2 Promotes Replication Fork Restart. *CellReports*, *7*(4), 1048–1055. https://doi.org/10.1016/j.celrep.2014.04.007

56. Pitsikas P, Lee D, Rainbow AJ (May 2007). "Reduced host cell reactivation of oxidative DNA damage in human cells deficient in the mismatch repair gene hMSH2". Mutagenesis. 22 (3): 235–43. doi:10.1093/mutage/gem008. PMID 17351251.

57. Quinet, A., Lemacon, D., Vindigni, A., Replication Fork Reversal: Players and Guardians. Review Article. Mol Cell. 2017 Dec 7;68(5):830-833. doi: 10.1016/j.molcel.2017.11.022.

58. Reid, E. (1998). Emery's Elements of Medical Genetics. *Journal of Medical Genetics*, *35*(9), 792–792. https://doi.org/10.1136/jmg.35.9.792-a

59. Sakofsky, C. J., Ayyar, S., & Malkova, A. (2012). Break-induced replication and genome stability. *Biomolecules*, *2*(4), 483–504. https://doi.org/10.3390/biom2040483

60. Schmidt, M. H. M., & Pearson, C. E. (2016). Disease-associated repeat instability and mismatch repair. *DNA Repair*. https://doi.org/10.1016/j.dnarep.2015.11.008

61. Seriola, A., Spits, C., Simard, J. P., Hilven, P., Haentjens, P., Pearson, C. E., & Sermon, K. (2011). Huntington's and myotonic dystrophy hESCs: Down-regulated trinucleotide repeat instability and mismatch repair machinery expression bupon differentiation. *Human Molecular Genetics*, *20*(1), 176–185. https://doi.org/10.1093/hmg/ddq456

62. Shah, S., Hile, S., & Eckert, K. (2010). Defective Mismatch Repair, Microsatellite Mutation Bias, and Variability in Clinical Cancer Phenotypes. *Cancer Research*, *70*(2), 431–435.

63. Sotiriou et al., 2016, Molecular Cell 64, 1127–1134 December 15, 2016 ª 201The Author(s). Published by Elsevier Inc. http://dx.doi.org/10.1016/j.molcel.2016.10.038

64. Srivastava, M., & Raghavan, S. C. (2015). Review DNA Double-Strand Break Repair Inhibitors as Cancer Therapeutics. *Chemistry & Biology*, *22*(1), 17–29. https://doi.org/10.1016/j.chembiol.2014.11.013

65. Sykora, P., Misiak, M., & Bohr, V. (2015). DNA polymerase β deficiency leads to neurodegeneration and exacerbates Alzheimer disease phenotypes. *Nucleic Acids Research*, *43*(3), 943–959.

66. Te´ cher et al., 2016, Cell Reports 14, 1114–1127 February 9, 2016 ª2016 The Authors http://dx.doi.org/10.1016/j.celrep.2015.12.093

67. Team, S. (2011). Signaling Pathway Integrated Knowledge Engine (SPIKE). Retrieved from https://www.cs.tau.ac.il/~spike/maps/spike00009.html

68. Tomasetti, C., & Vogelstein, B. (2015). Variation in cancer risk among tissues can be explained by the number of stem cell divisions. *Science*, *347*(6217), 78–81. https://doi.org/10.1126/science.1260825

69. Wang, G., & Vasquez, K. (2009). Models for chromosomal replication-independent non-B DNA structure-induced genetic instability. *NIH Public Access*, *48*(4), 286–298.

70. Wang, G., & Vasquez, K. M. (2009). Models for chromosomal replication-independent Non-B DNA structure-induced genetic instability. *Molecular Carcinogenesis*. https://doi.org/10.1002/mc.20508

71. Wells, R. D., Parniewski, P., Pluciennik, A., Bacolla, A., Gellibolian, R., & Jaworski, A. (1998). Small slipped register genetic instabilities in Escherichia coli in triplet repeat sequences associated with hereditary neurological diseases. *Journal of Biological Chemistry*, *273*(31), 19532–19541. https://doi.org/10.1074/jbc.273.31.19532

72. Wind N, Dekker M, Berns A, Radman M, te Riele H (July 1995). "Inactivation of the mouse Msh2 gene results in mismatch repair deficiency, methylation tolerance, hyperrecombination, and predisposition to cancer". Cell. 82 (2): 321–30. doi:10.1016/0092-8674(95)90319-4. PMID 7628020.

73. Wyatt, H. D. M., Sarbajna, S., Matos, J., & West, S. C. (2013). Article Coordinated Actions of SLX1-SLX4 and MUS81-EME1 for Holliday Junction Resolution in Human Cells. *Molecular Cell*, *52*(2), 234–247. https://doi.org/10.1016/j.molcel.2013.08.035

74. Xing, M., Xiaohui, W., Palmai-Pallag, T., Shen, H., Helleday, T., Hickson, I., & Ying, S. (2015). Acute MUS81 depletion leads to replication fork slowing and a constitutive DNA damage response. *Oncotarget*, *6*(35).

75. Zhao, J., Bacolla, A., Wang, G., & Vasquez, K. M. (2010). Non-B DNA structure-induced genetic instability and evolution. *Cellular and Molecular Life Sciences*. https://doi.org/10.1007/s00018-009-0131-2

76. Zhou, X., Delucia, M., & Ahn, J. (2016). SLX4-SLX1 Protein-independent Down-regulation of MUS81-EME1 Protein by HIV-1 Viral Protein R ( Vpr ) *, *291*(33), 16936–16947. https://doi.org/10.1074/jbc.M116.721183

77. Zimmer, J., Badie, S., Niedzwiedz, W., Lai, X., Broderick, R., & Tarsounas, M. (2017). stress tolerance and chromosome segregation in, (May). https://doi.org/10.1038/ncomms15983